Fractal Dimensions in Theory and Practice

A thesis submitted to attain the degree of

MASTER OF SCIENCES

(MSc ETH Zürich Mathematics)

presented by

Nick Kunz

supervised by

Dr. Sara Kališnik Hintz (ETH Zürich)

May 18, 2024

ETH zürich

Abstract

Fractals are geometric shapes that show arbitrary fine structures and self-similarities. The dimension of a space indicates how its size changes when we scale the space. Due to their unique scaling properties, fractals can display non-integer dimensions. In this thesis we characterise fractals and analyse different fractal dimensions including some standard fractal dimensions like the Minkowski and Hausdorff dimension. We construct the Vietoris-Rips complex for finite metric spaces which leads to the definition of persistent homology and persistent homology dimension. Furthermore, we define magnitude for metric spaces and explore how magnitude changes when scaling a space revealing the magnitude dimension. We develop methods to estimate persistent homology and magnitude dimensions, which are then compared to the Minkowski and Hausdorff dimensions.

Acknowledgements

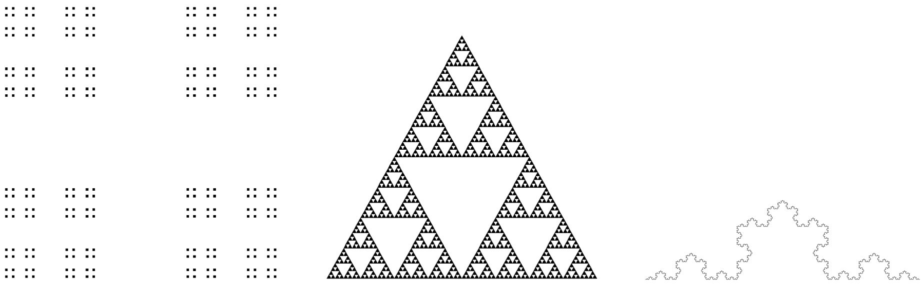
I would like to express my gratitude to my advisor Dr. Sara Kališnik Hintz for her guidance and support. Her expertise and constructive feedback have played a central role in shaping this thesis.
I am also immensely thankful to Dr. Bastian Rieck and Katharina Limbeck for their generous assistance with the coding aspects of this thesis.

Contents

Chapter 1

Introduction

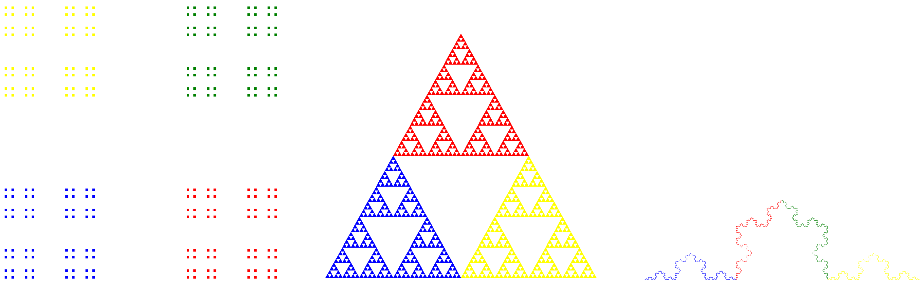
Fractals are geometric objects which display detailed structures at arbitrary fine scale. When zooming in on a fractal, it reveals its infinitely fine structure. Some famous examples of fractals include the Cantor dust, the Sierpinski triangle and the Koch Curve shown in Figure 1.1.



(a) Cantor dust (b) Sierpinski triangle (c) Koch curve

Figure 1.1: Three examples of fractals.

Another characteristic that fractals share is their self-similarity. Often, a fractal can be decomposed into several copies of itself at lower scales. The self-similarities of the Cantor dust, the Sierpinski triangle and the Koch Curve become apparent when colouring each copy in its decomposition with a different colour (see Figure 1.2). The Cantor dust is made up of 4 copies of itself scaled by $\frac{1}{3}$, the Sierpinski triangle of 3 copies scaled by $\frac{1}{2}$, and the Koch curve of 4 copies scaled by $\frac{1}{3}$.
This ties into another characteristic of fractals; their unique scaling property. If we scale a line segment, a square or a cube by a factor of 2, their respective volumes increase 2-, 4- and 8-fold. This is precisely 2 to the power of their respective dimension. Fractals behave somewhat differently in that regard. If we scale the Sierpinski triangle by a factor of 2, by the self-similarity described above, we end up with 3 copies of the original. So we can say its 'size' has increased 3-fold. This is a non-integer power of the scaling factor 2. By analogy to the above example, we expect the Sierpinski triangle's 'dimension' to be $\frac{\log(3)}{\log(2)} \approx 1.585$. We formalise this intuitive idea of 'fractured' dimension in Sections 4.1 and 4.2, where we introduce the Minkowski and Hausdorff dimen-



(a) Cantor dust (b) Sierpinski triangle (c) Koch curve

Figure 1.2: Visualising the self-similarity of fractals.

sions. For example, the Minkowski dimension of a compact metric space (X, d) is defined as

$$\dim_{\text{mink}}(X) := \lim_{\varepsilon \to 0^+} \frac{\log(N(X, \varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}$$

where $N(X, \varepsilon)$ denotes the ε-covering number of X, that is, the minimum number of ε-balls needed to cover X.

A different fractal dimension can be defined by means of persistent homology. Persistent homology is a tool from algebraic topology that captures the topological features of a space and shows how these features persist over different scales. In Section 5.1 we introduce the Vietoris-Rips complex for metric spaces. The k-simplices of the Vietoris-Rips complex at some scale $r \geq 0$ consist of all $(k+1)$-point sets which have diameter less or equal to r. By varying $r \in \mathbb{R}_{\geq 0}$, we get a sequence of simplicial complexes with inclusion maps between them, called a filtration. Figure 1.3 depicts Vietoris-Rips complexes at eight different scales of a metric space obtained by sampling 20 points from an annulus.

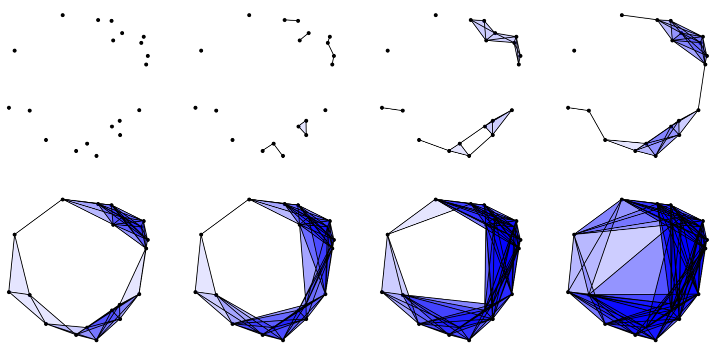
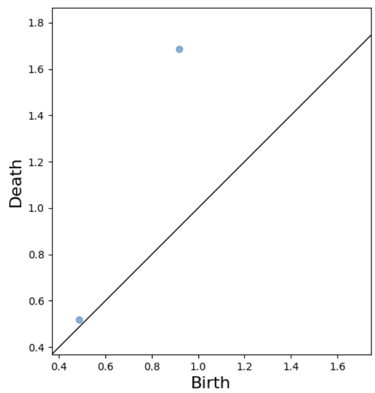


Figure 1.3: Example of Vietoris-Rips complexes.

By applying the homology functor to the Vietoris-Rips filtration, we obtain a persistence module. Persistence modules that arise from the Vietoris-Rips complex of finite spaces are decomposable into so-call interval modules. Each interval module $U(b, d)$ represents a topological feature which appears at scale b and disappears at scale d. We can visualise the multiset of lifetime intervals $[b, d]$ using barcodes and persistence diagrams. A barcode plots the lifetime intervals, while a persistence diagram plots the tuples (b, d). Figure 1.4 depicts the barcode and persistence diagram of the 1st persistence module PH_1, for the 20 points sampled from an annulus. In this example, PH_1 has two generators: one is only visible in the third Vietoris-Rips complex in 1.3, the other persists from the fifth to the seventh and is then filled in.

(a) Barcode



(b) Persistence diagram

Figure 1.4: Visualisation of PH_1 by a barcode and a persistence diagram.

In Section 5.1 we show that in the case of finite metric spaces, there is a correspondence between the bounded intervals in PH_0 and the edges of a minimal spanning trees on the space. Using this relation, we are able to generalise the definition of topological dimensions which involve minimal spanning trees. In Section 5.2.1 we expand on the idea of analysing minimal spanning trees on extremal subsets. We define a persistent homology dimension for compact metric spaces (X, d) as

$$\dim_{\mathrm{PH}}^k(X) := \inf\{\alpha \geq 0 \colon \exists C \in \mathbb{R} \colon \forall A \subseteq X \text{ finite} \colon E_\alpha^k(A) \leq X\},$$

where

$$E_\alpha^k(A) := \sum_{(b,d) \in \mathrm{PH}_k(A)} (d - b)^\alpha.$$

In Section 5.2.2 we generalise the notion of analysing minimal spanning trees on random subsets to analysing persistence modules on random subsets. We define a persistent homology dimension for probability measures μ on a compact metric space. For $X_n \subseteq X$ i.i.d. samples according to μ, we define

$$\dim_{\mathrm{PH}}^k(\mu) := \inf\left\{d > 0 \colon \exists C \in \mathbb{R} \colon \lim_{n\to\infty} \mathbb{P}\left[L_k(X_n) \leq C n^{\frac{d-1}{d}}\right] = 1\right\},$$

where

$$L_k(X_n) := \sum_{(b-d) \in \mathrm{PH}_k(X_n)} (d - b).$$

The latter definition has the advantage of being easy to approximate. We expect the asymptotic slope of $L_k(X_n)$ over n in a log-log plot to be close to $\frac{d-1}{d}$ for $d = \dim_{\mathrm{PH}}^k(\mu)$. Figure 1.5 shows the log-log plot of $L_0(X_n)$ for samples of the Sierpinski triangle according to a uniform distribution μ. The asymptotic slope

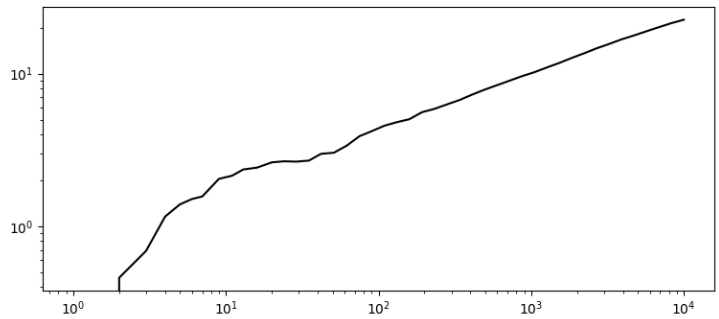


Figure 1.5: L_0 for uniform samples on the Sierpinski triangle.

is approximately 0.3187 which means that

$$\dim_{\mathrm{PH}}^k(\mu) \approx 1.4679.$$

This is close to the Minkowski dimension of the Sierpinski triangle, which is known to be

$$\dim_{\mathrm{mink}}(\mathcal{S}) = \frac{\log(3)}{\log(2)} \approx 1.5850.$$

Another fractal dimension of a metric spaces is the magnitude dimension. Magnitude is an invariant, which, for finite spaces, can be thought of as a measure of the 'effective number of points'. If some points lie very close together, they become indistinguishable from one another and appear 'effectively' as one point. Consider for example the 3-point space in Figure 1.6 viewed at different scales. When viewed at a large scale, all three points can be easily be distinguished. At medium scale, the left point can be distinguished from the two points on the right. But the two points on the right cannot be distinguished and appear as one. If we zoom out all the way, viewing the space at a small scale, the three points are indistinguishable and appear as one single point.

(a) Large scale (b) Medium scale (c) Small scale

Figure 1.6: A three point space viewed at different scales.

The magnitude function of a metric space assigns to each scale $t > 0$ the magnitude of the scaled metric space tX, which has the same points as X but its metric is scaled by t. The rate at which the magnitude function changes encodes information about the 'instantaneous dimension' of the metric space. The instantaneous magnitude dimension is defined as the growth rate of the magnitude function, that is, the slope of the magnitude function when plotted on a log-log scale,

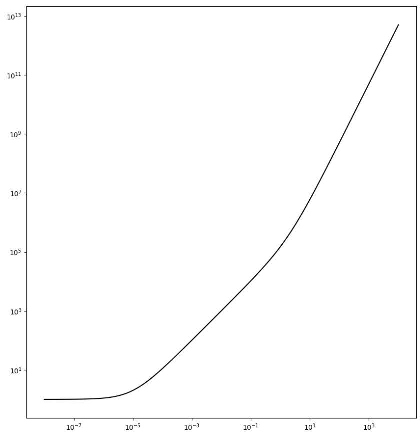
$$\dim_{\text{mag}}^{\text{inst}}(X,t) := \frac{\mathrm{d}\log(|sX|)}{\mathrm{d}\log(s)}\bigg|_{s=t} = \frac{s}{|sX|}\frac{\mathrm{d}|sX|}{\mathrm{d}s}\bigg|_{s=t}.$$

Consider a long, thin rectangle and view it at different scales as shown in Figure 1.7. At a large scale, we can see that the rectangle is 2-dimensional. When we zoom out and view it at a medium scale, the rectangle appears as a 1-dimensional line. If we zoom out even more, and view it at a small scale, the rectangle appears as a single point, which is 0 dimensional. Figure 1.8a shows

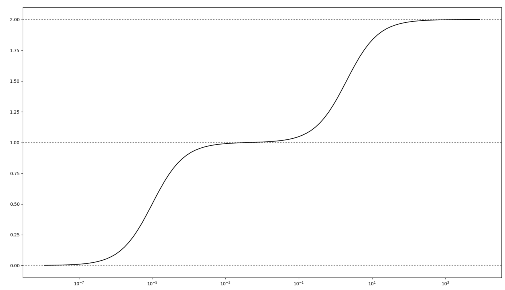
(a) Large scale. (b) Medium scale. (c) Small scale.

Figure 1.7: A long, thin rectangle at different scales.

the log-log plot of the magnitude function over the scale parameter. Figure 1.8b shows the instantaneous magnitude dimension of the long, thin rectangle. We can see that it has instantaneous magnitude dimension 0 at small scale, 1 at medium scale and 2 at large scale.



(a) Magnitude function.



(b) Instantaneous growth rate.

Figure 1.8: Magnitude function and instantaneous growth rate.

The magnitude dimension of a metric space is defined as the asymptotic growth rate of the magnitude function

$$\dim_{\text{mag}}(X) := \lim_{t\to\infty}\frac{\log(|tX|)}{\log(t)} = \lim_{t\to\infty}\dim_{\text{mag}}^{\text{inst}}(X,t).$$

We can see from Figure 1.8b that the magnitude dimension of the long, thin rectangle is 2.

This thesis is dedicated to the characterisation of fractals and the comparison of various fractal dimensions. We start by establishing the foundational concepts for metric spaces in Chapter 2. In Chapter 3, we characterise fractals and provide examples. We also introduce three standard fractal dimensions in Chapter 4: The Minkowski, the Hausdorff, and the minimal spanning tree dimensions. Chapter 5 is dedicated to defining persistent homology through the Vietoris-Rips complex and introducing two fractal dimensions based on persistent homology. In Chapter 6 we present the concept of magnitude. We analyse its continuity properties and study the magnitude function. This leads us to define another fractal dimension – the magnitude dimension. In Section 6.4 we study spread, a measurement of size for metric spaces similar to the magnitude, and we define spread dimension. Lastly, in Chapter 7, we estimate the persistent homology and magnitude dimension of several fractals and compare them to their Minkowski and Hausdorff dimension.

Chapter 2

Metric Spaces

This section covers some of the basic definitions for metric spaces. We define metric and pseudometric spaces, and we introduce the Hausdorff and Gromov-Hausdorff distance as a measure of how much two metric spaces differ. We refer to [13] for a more detailed introduction to metric spaces.

Definition 2.1. *A* **metric space** (X, d) *is a set* X *together with a map*

$$d\colon X \times X \to \mathbb{R}$$

called the **metric** *or* **distance** *which satisfies the following:*

- *Positivity:*

$$\forall x, y \in X\colon d(x, y) \geq 0 \wedge (d(x, y) = 0 \iff x = y)\,.$$

- *Symmetry:*

$$\forall x, y \in X\colon d(x, y) = d(y, x).$$

- *Triangle inequality:*

$$\forall x, y, z \in X\colon d(x, z) \leq d(x, y) + d(y, z).$$

In the definition of a **pseudometric space***, the positivity condition is relaxed to the weaker condition*

$$\forall x, y \in X\colon d(x, y) \geq 0 \wedge d(x, x) = 0.$$

Example 2.1. *The Euclidean metric on* $\mathbb{R}^n$ *is given by*

$$d(x, y) \coloneqq \left(\sum_{i=1}^{n} (y_i - x_i)^2 \right)^{\frac{1}{2}}, \quad \forall x, y \in \mathbb{R}^n.$$

For the entirety of this thesis we assume that all metric spaces and their subspaces are non-empty.

Definition 2.2. *The* **diameter** *of a metric space* (X, d) *is defined as*

$$\operatorname{diam}(X) \coloneqq \sup\{d(a, b)\colon a, b \in X\}.$$

2.1 Gromov-Hausdorff Distance

When analysing metric spaces, it is useful to measure how much they differ from each other. We can measure the distance between two subspaces of a metric space using the Hausdorff distance. The Hausdorff distance between two subspaces is the maximum distance a point of one subspace has to the other subspace. That means, two subspaces are close to each other if any point of one subspace is close to some point in the other subspace.

Definition 2.3. *Let (X, d^X) be a metric space. The* **Hausdorff distance** d_{H}^X *on the set of non-empty subspaces of X is defined by*

$$d_{\mathrm{H}}^X(A, B) := \max\left\{\sup_{a\in A} d^X(a, B), \sup_{b\in B} d^X(A, b)\right\},$$

where

$$d^X(a, B) := \inf_{b\in B} d^X(a, b), \quad d^X(A, b) := \inf_{a\in A} d^X(a, b).$$

Example 2.2. *Consider the Euclidean space $X = \mathbb{R}^2$ and two subspace $A \subseteq X$ consisting of 10 points and $B \subseteq \mathbb{R}^2$ consisting of 5 points. Figure 2.1 shows the two subspaces and where the Hausdorff distance between them is attained.*

(a) A (red) and B (blue).

(b) $d_{\mathrm{H}}^X(A, B)$.

Figure 2.1: Hausdorff distance between two metric subspaces.

Proposition 2.1. *The Hausdorff distance defines a metric on the set of non-empty, compact metric subspaces of X.*

Proof. If $A, B \subseteq X$ are non-empty, compact metric subspaces, then the suprema and infima in the definition of $d_{\mathrm{H}}^X(A, B)$ are attained. So we can write

$$d_{\mathrm{H}}^X(A, B) = \max\left\{\max_{a\in A} d^X(a, B), \max_{b\in B} d^X(A, b)\right\},$$

where

$$d^X(a, B) = \min_{b\in B} d^X(a, b), \quad d^X(A, b) = \min_{a\in A} d^X(a, b).$$

Hence $d_{\mathrm{H}}^X(A, B) < \infty$, which means that d_{H}^X is well-defined.

- Positivity:
  By the positivity of d^X, we know that $d_{\mathrm{H}}^X(A, B) \geq 0$. It is also clear that $d_{\mathrm{H}}^X(A, A) = 0$. Furthermore, if $d_{\mathrm{H}}^X(A, B) = 0$, then $\max_{a\in A} d^X(a, B) = 0$, so for all $a \in A$,

$$d^X(a, B) = \min_{b\in B} d^X(a, b) = 0.$$

Therefore, for all $a \in A$, there exists $b \in B$ such that $d^X(a,b) = 0$. By the positivity of d^X, this implies that $b = a$. Thus $A \subseteq B$. Similarly, one finds that $B \subseteq A$, which implies $A = B$.

- Symmetry:
  It follows immediately from the definition that $d_{\mathrm{H}}^X(A,B) = d_{\mathrm{H}}^X(B,A)$.

- Triangle inequality:
  Let $C \subseteq X$ be a non-empty, compact metric subspace and fix $a \in A$. Since B is compact, there exists $b \in B$ such that $d^X(a,B) = d^X(a,b)$. Then

$$\begin{aligned} d^X(a,C) &= \min_{c \in C} d^X(a,c) \\ &\leq \min_{c \in C} d^X(a,b) + d^X(b,c) \\ &= d^X(a,B) + d^X(b,C) \\ &\leq \max_{a' \in A} d^X(a',B) + \max_{b' \in B} d^X(b',C) \\ &\leq d_{\mathrm{H}}^X(A,B) + d_{\mathrm{H}}^X(B,C). \end{aligned}$$

  Since this holds for all $a \in A$, we find that

$$\max_{a \in A} d^X(a,C) \leq d_{\mathrm{H}}^X(A,B) + d_{\mathrm{H}}^X(B,C).$$

  Switching the roles of A and C, we find that

$$\max_{c \in C} d^X(A,c) \leq d_{\mathrm{H}}^X(A,B) + d_{\mathrm{H}}^X(B,C).$$

  Therefore

$$d_{\mathrm{H}}^X(A,C) \leq d_{\mathrm{H}}^X(A,B) + d_{\mathrm{H}}^X(B,C).$$

□

We are able to measure how much two metric spaces differ by embedding them both in a common metric space, and considering the Hausdorff distance. The Gromov-Hausdorff distance is the infimum of the Hausdorff dimension over all possible embeddings.

Definition 2.4. *The* **Gromov–Hausdorff distance** *between two non-empty metric spaces* (A, d^A) *and* (B, d^B) *is defined as*

$$d_{\mathrm{GH}}(A,B) := \inf_{X, \varphi_X, \psi_X} d_{\mathrm{H}}^X(\varphi_X(A), \psi_X(B)),$$

where the infimum is taken over all metric spaces (X, d^X) *and isometric embeddings* $\varphi_X \colon A \hookrightarrow X$ *and* $\psi_X \colon B \hookrightarrow X$*.*

Example 2.3. *Figure 2.2 depicts the embedding of two metric spaces into a common metric space that minimises their Hausdorff distance.*

Remark 2.1. *The Gromov-Hausdorff distance turns the set of non-empty, compact metric spaces into a pseudometric space, where* A *and* B *have Gromov-Hausdorff distance* 0 *if and only if they are isometric.*

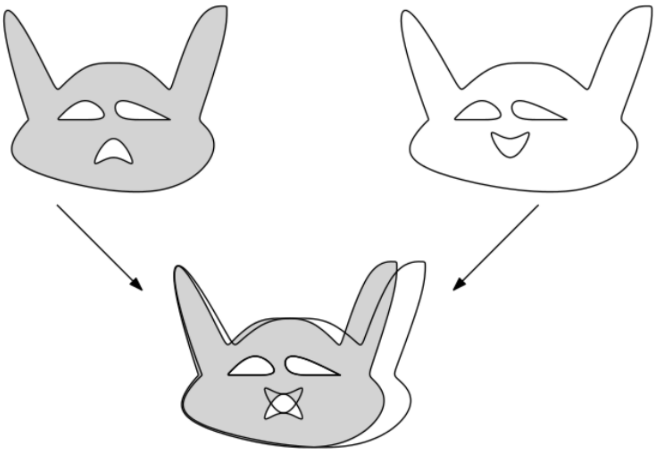


Figure 2.2: Visualising the Gromov-Hausdorff distance (Figure 3 in [8]).

2.2 Minimal Spanning Trees

In this section we introduce minimal spanning trees and present Kruskal's algorithm for finding minimal spanning trees. We refer to [4] for more details.

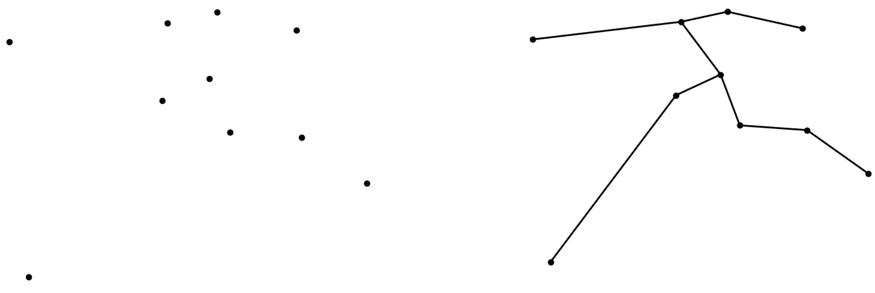
Definition 2.5. *Let (X, d) be a finite metric space. A* **spanning tree** *on X is a graph $T(X)$ which connects all the points in X without any cycles. The* **weight** *$\|e\|$ of an edge e between two points $x, y \in X$ is given by the distance between the two points*

$$\|e\| := d(x, y).$$

A **Minimal spanning tree (MST)** *on (X, d) is a spanning tree $T(X)$ which minimises the total edge weight*

$$\sum_{e \in T(X)} \|e\|.$$

Example 2.4. *Figure 2.3 depicts a metric space of* 10 *points in* $\mathbb{R}^2$ *and a MST on it.*



(a) Metric space. (b) MST.

Figure 2.3: Metric space with MST.

One algorithm for finding a MST is Kruskal's algorithm, where we start with no edges between the vertices and add the shortest edge which does not create a cycle until we end up with a spanning tree. We refer to [4] for further details on minimal spanning trees and Kruskal's algorithm.

Chapter 3

Fractals

The goal of this chapter is to present the characteristics which describe fractals and give some examples. We construct the Cantor set, the Sierpinski triangle and the Koch curve and snowflake. The reader is referred to [7] and [6] for further reading.

3.1 Fractals: Characteristics and Examples

The term ***fractal*** refers to a certain class of geometric objects. While lacking a universal definition, there are three commonly agreed-upon characteristics that describe a fractal.

- Fractals have the capacity to exhibit detailed structure at arbitrary small scales, revealing its infinitely fine structure when zooming in on it.
- A fractal usually exhibits some kind of self-similarity, where intricate patterns repeat themselves at different scales within the structure. Often, this can be seen in the composition of a fractal, with each part resembling a scaled-down copy of the whole.
- Fractals have distinctive scaling properties and 'fractured' dimension. When scaling a line segment, a square or a cube by a factor of 2, their respective volume increase 2-, 4- and 8-fold respectively. This is precisely 2 to the power of their respective dimension. For fractals, this exponent can take on non-integer values.

Example 3.1. *The* **Cantor set** $\mathcal{C}$ *is a fractal. It can be constructed as follows: start with the closed interval* $C_0 := [0, 1]$*. Remove the middle third of the interval to get* C_1*, that is*

$$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right].$$

Next, remove the middle third of each of the two intervals in C_1 *to get*

$$C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right].$$

Continue this process iteratively, obtaining C_n *by removing the middle thirds from all intervals* C_{n-1} *is made up of. Then* C_n *is the the disjoint union of* 2^n

closed intervals of length 3^{-n}. The Cantor set $\mathcal{C}$ consists exactly of the points that remain after continuing this process infinitely, that is

$$\mathcal{C} := \bigcap_{n=0}^{\infty} C_n.$$

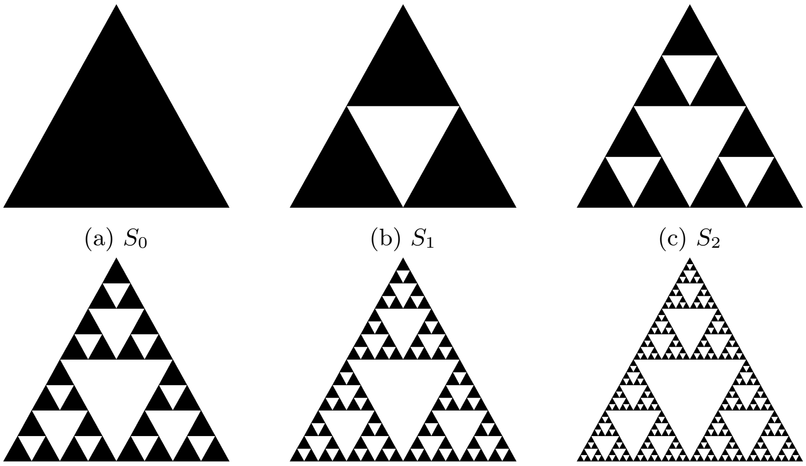
Figure 3.1 shows the first six sets of intervals $C_0, \ldots, C_5$ in the construction of the Cantor set.

Figure 3.1: Construction of the Cantor set.

Example 3.2. *Another example of a fractal is the* **Sierpinski triangle** *$\mathcal{S}$. It can be constructed as follows: start with an equilateral triangle S_0. Now subdivide the triangle into four smaller equilateral triangles and remove the middle one to get S_1. Continue this process for each of the remaining triangles. We obtain S_n from S_{n-1} by subdividing all triangles into four smaller equilateral triangles and removing the middle ones. The Sierpinski triangle is the limit of this construction*

$$\mathcal{S} := \bigcap_{n=0}^{\infty} S_n.$$

Figure 3.2 depicts $S_0, \ldots, S_5$ from the construction of the Sierpinski triangle.

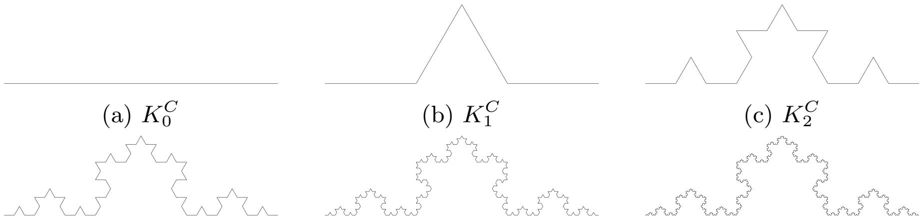


(a) S_0 (b) S_1 (c) S_2

(d) S_3 (e) S_4 (f) S_5

Figure 3.2: Construction of the Sierpinski triangle.

Example 3.3. *The* **Koch curve** *$\mathcal{K}^C$ and the* **Koch snowflake** *$\mathcal{K}^S$ are fractals. They can be constructed as follows: start with a line segment K_0^C. Now divide* *the line segment into three equal parts. Draw an equilateral triangle that has the middle segment as a base and is pointed upwards, and remove the base. This yields K_1^C consisting of four equal line segments. Repeat this process on each of the line segments. The Koch curve $\mathcal{K}^C$ is the limit of this process. Figure 3.3 shows $K_0^C, \ldots, K_5^C$ from this construction. The Koch snowflake $\mathcal{K}^S$ is obtained*



(a) K_0^C (b) K_1^C (c) K_2^C

(d) K_3^C (e) K_4^C (f) K_5^C

Figure 3.3: Construction of the Koch curve.

by replacing the three sides of an equilateral triangle with three Koch curves pointing outwards (see Figure 3.4).

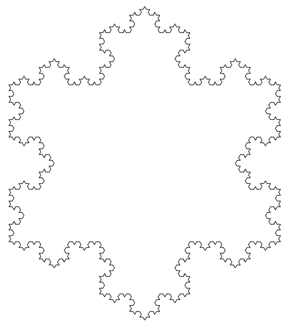


Figure 3.4: Koch snowflake.

Chapter 4

Standard Fractal Dimensions

The aim of this chapter is to present three standard fractal dimensions. In Section 4.1 we define the Minkowski dimension for compact metric spaces and compute it for the Cantor set. We define the Hausdorff dimension in Section 4.2 and compute the Hausdorff dimension of the Cantor set. We discuss a result that simplifies the computation of the Hausdorff dimension is case of self similar fractals and show that in that case the Hausdorff and Minkowski dimension agree (Theorem 4.1). For this section we refer to [7]. As preparation for Chapter 5, we define the MST dimension in Section 4.3 and prove that it is equal to the Minkowski dimension (Theorem 4.2), following [9].

4.1 Minkowski Dimension

The 'dimension' of a metric space measures how the 'size' of the metric space changes as we scale the space. Consider, for example, a line segment, a square and a cube, and scale them by a factor of 2, as depicted in Figure 4.1. Then their 'sizes' increases by a factor of 2, 4 or 8 respectively, which is 2 to the power of their respective definition. To make this notion precise, we introduce the covering and the packing number as two measurements of the size of a metric space. Both of these measurements lead to the definition of the Minkowski dimension, which is the asymptotic growth rate of the covering or the packing number.
For the remainder of this section, let (X, d) be a compact metric space.

Definition 4.1. *Let $\varepsilon > 0$. The ε-**covering number** $N(X, \varepsilon)$ is the minimum number of open ε-balls needed to cover X. The ε-**packing number** $M(X, \varepsilon)$ is the maximum number of disjoint open ε-balls in X.*

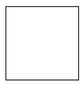
Example 4.1. *Figure 4.2 depicts a minimal ε-cover and a maximal δ-packing of a square.*

Lemma 4.1. *For all $\varepsilon > 0$ we have*

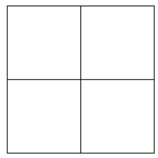
$$N(X, 2\varepsilon) \leq M(X, \varepsilon) \leq N(X, \varepsilon).$$

(a) Line segment at scale 1.

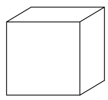
(b) Line segment at scale 2.



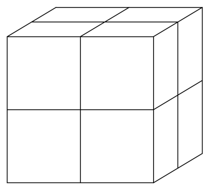
(c) Square at scale 1.



(d) Square at scale 2.



(e) Cube at scale 1.



(f) Cube at scale 2.

Figure 4.1: A line segment, a square and a cube at different scales.

Proof. Write $M := M(X, \varepsilon)$ and let $x_1, \ldots, x_M \in X$ be such that $\{B(x_i, \varepsilon)\}_i$ is a maximal ε-packing of X.

- $N(X, 2\varepsilon) \leq M(X, \varepsilon)$:
  By the maximality of $\{B(x_i, \varepsilon)\}_i$, ever $y \in X$ must satisfy $d(y, x_i) < 2\varepsilon$ for some i, otherwise we could add $B(y, \varepsilon)$ to get a larger ε-packing. Therefore, $\{B(x_i, 2\varepsilon)\}_i$ is a (2ε)-covering of X.

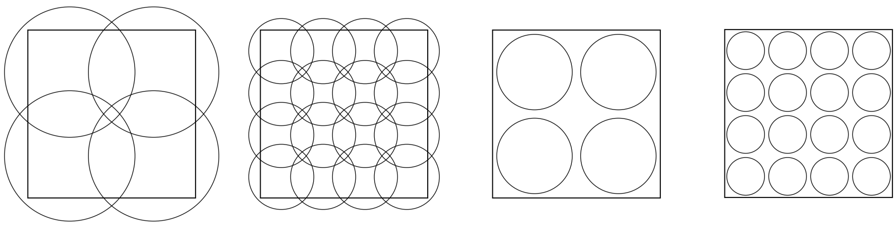
- $M(X, \varepsilon) \leq N(X, \varepsilon)$:
  Since $\{B(x_i, \varepsilon)\}_i$ are disjoint, we find $d(x_i, x_j) \geq 2\varepsilon$ whenever $i \neq j$. This means that in any ε-covering of X, all x_i must lie in different ε-balls. Hence every covering must consist of at least M ε-balls.

□

Consider again the ε-covering and δ-packing from Example 4.1. By halving ε and δ, the covering and packing number increase 4-fold (see Figure 4.3). This is precisely 2 to the power of the 'dimension' of the square. We use this intuition as a definition of the Minkowski dimension as the asymptotic growth rate of the ε-covering number as $\varepsilon \to 0^+$.

(a) ε-covering. (b) δ-packing.

Figure 4.2: An ε-covering and a δ-packing of a square.



(a) ε-covering. (b) $\frac{\varepsilon}{2}$-covering. (c) δ-packing. (d) $\frac{\delta}{2}$-packing.

Figure 4.3: Coverings and packings of a square.

Definition 4.2. *The* **lower Minkowski dimension** *of X is*

$$\underline{\dim}_{\text{mink}}(X) \coloneqq \liminf_{\varepsilon \to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}$$

and the **upper Minkowski dimension** *of X is*

$$\overline{\dim}_{\text{mink}}(X) \coloneqq \limsup_{\varepsilon \to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}.$$

If the lower and upper Minkowski dimension coincide, then the **Minkowski dimension** *exist and is equal to this limit*

$$\dim_{\text{mink}}(X) \coloneqq \lim_{\varepsilon \to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}.$$

The Minkowski dimensions is also known as the **box dimension** *or* **box-counting dimension**.

Proposition 4.1. *The lower and upper Minkowski dimensions can also be expressed in terms of the packing number:*

$$\underline{\dim}_{\text{mink}}(X) = \liminf_{\varepsilon \to 0^+} \frac{\log(M(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)},$$

$$\overline{\dim}_{\text{mink}}(X) = \limsup_{\varepsilon \to 0^+} \frac{\log(M(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}.$$

Hence, if the Minkowski dimension exists, then

$$\dim_{\text{mink}}(X) = \lim_{\varepsilon \to 0^+} \frac{\log(M(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}.$$

Proof. For all $\varepsilon > 0$, by Lemma 4.1 and the monotonicity of log we find that

$$\begin{aligned}\liminf_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} &\geq \liminf_{\varepsilon\to 0^+} \frac{\log(M(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \\ &\geq \liminf_{\varepsilon\to 0^+} \frac{\log(N(X,2\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \\ &= \liminf_{\varepsilon\to 0^+} \frac{\log(N(X,2\varepsilon))}{\log(2)+\log\left(\frac{1}{2\varepsilon}\right)} \\ &= \liminf_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log(2)+\log\left(\frac{1}{\varepsilon}\right)}.\end{aligned}$$

Since $\log\left(\frac{1}{\varepsilon}\right) \to \infty$ as $\varepsilon \to 0^+$, we have

$$\liminf_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log(2)+\log\left(\frac{1}{\varepsilon}\right)} = \liminf_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}.$$

Therefore

$$\liminf_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \geq \liminf_{\varepsilon\to 0^+} \frac{\log(M(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \geq \liminf_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)},$$

which proves equality

$$\liminf_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} = \liminf_{\varepsilon\to 0^+} \frac{\log(M(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}.$$

Analogously, we find that

$$\limsup_{\varepsilon\to 0^+} \frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} = \limsup_{\varepsilon\to 0^+} \frac{\log(M(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)}.$$

□

Example 4.2. *Let us show that the Minkowski dimension of the Cantor set* $\mathcal{C}$ *exists and is given by*

$$\dim_{\text{mink}}(\mathcal{C}) = \frac{\log(2)}{\log(3)}.$$

We prove the following claim in preparation.

Lemma 4.2. *Let* $n \in \mathbb{N}$. *The* (3^{-n})*-covering number of the Cantor set is*

$$N(C, 3^{-n}) = 2^{n+1}.$$

Proof. Consider C_n from the construction of the Cantor set (see Example 3.1). C_n consists of 2^n intervals of length 3^{-n}. The endpoints of each interval are also points in $\mathcal{C}$ and have a distance of at least 3^{-n} between them. Therefore these 2^{n+1} endpoints lie in different 3^{-n}-balls and so

$$N(\mathcal{C}, 3^{-n}) \geq 2^{n+1}.$$

On the other hand, the 2^{n+1} (3^{-n})-balls centered at the endpoints of all interval in C_n form a cover of C_n and hence $\mathcal{C}$. Thus

$$N(\mathcal{C}, 3^{-n}) \leq 2^{n+1},$$

proving the claim. □

By Lemma 4.2 and the monotonicity of the covering number, we find that

$$2^n \leq N(\mathcal{C}, \varepsilon) \leq 2^{n+1}$$

whenever

$$3^{-n} \leq \varepsilon < 3^{-(n-1)}.$$

Therefore

$$\frac{\log(2^n)}{\log(3^n)} \leq \frac{\log(N(\mathcal{C}, \varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \leq \frac{\log(2^{n+1})}{\log(3^{n-1})},$$

which simplifies to

$$\frac{\log(2)}{\log(3)} \leq \frac{\log(N(\mathcal{C}, \varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \leq \frac{n+1}{n-1}\frac{\log(2)}{\log(3)}.$$

Therefore

$$\underline{\dim}_{\text{mink}}(\mathcal{C}) = \liminf_{\varepsilon \to 0^+} \frac{\log(N(\mathcal{C}, \varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \geq \frac{\log(2)}{\log(3)}.$$

And

$$\overline{\dim}_{\text{mink}}(\mathcal{C}) = \limsup_{\varepsilon \to 0^+} \frac{\log(N(\mathcal{C}, \varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} \leq \lim_{n \to \infty} \frac{n+1}{n-1}\frac{\log(2)}{\log(3)} = \frac{\log(2)}{\log(3)}.$$

So the lower and upper Minkowski dimension must agree and are equal to

$$\dim_{\text{mink}}(\mathcal{C}) = \frac{\log(2)}{\log(3)}.$$

4.2 Hausdorff Dimension

The Hausdorff dimension takes a slightly different approach in measuring the change in 'size' of a metric space as we scale it. In contrast to the Minkowski dimension we consider coverings made up of arbitrary set that have diameter less or equal to δ instead of δ-balls. Since this approach is similar to the one taken Section 4.1, the Minkowski and Hausdorff dimensions agree on a large class of fractals (see Theorem 4.1).
For the entirety of this section, let (X, d) denote an arbitrary metric space.

Definition 4.3. *Let $A \subseteq X$ and $\delta > 0$. A* **δ-cover** *of A is a family $\{U_i\}_i$ of subsets $U_i \subseteq X$ such that* $\operatorname{diam}(U_i) \leq \delta$ *and*

$$A \subseteq \bigcup_i U_i.$$

If, in addition, $\{U_i\}_i$ is countable, we call it a **countable δ-cover** *of A.*

For $\delta > 0$ and $s \geq 0$ consider the quantity

$$\mathcal{H}^s_\delta(A) := \inf\left\{\sum_i \operatorname{diam}(U_i)^s \colon \{U_i\}_i \text{ is a countable } \delta\text{-cover of } A\right\},$$

defined on all subsets $A \subseteq X$. Every δ-cover of A is also a δ'-cover whenever $\delta' \geq \delta$. Therefore $\mathcal{H}^s_\delta(A) \geq \mathcal{H}^s_{\delta'}(A)$, which means that $\mathcal{H}^s_\delta(A)$ is monotonically non-increasing in δ. So the limit of $\mathcal{H}^s_\delta(A)$ as $\delta \to 0^+$ exists but might be infinite.

Definition 4.4. *The function*

$$\mathcal{H}^s\colon \mathcal{P}(X) \to [0,\infty],\ A \mapsto \mathcal{H}^s(A) := \lim_{\delta\to 0^+} \mathcal{H}^s_\delta(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A)$$

defined on the family $\mathcal{P}(X)$ *of all subsets of* X *is called the* s**-dimensional Hausdorff measure** *on* X.

As an outer metric measure, the Hausdorff measure is Borel.

Proposition 4.2. *Let* $A \subseteq X$ *be a Borel set and let* $0 \le s < t$. *Then*

$$\mathcal{H}^s(A) < \infty \implies \mathcal{H}^t(A) = 0,$$
$$\mathcal{H}^t(A) > 0 \implies \mathcal{H}^s(A) = \infty.$$

Proof. Let $\delta > 0$ and $\{U_i\}_i$ be a countable δ-cover of A. Then

$$\sum_i \operatorname{diam}(U_i)^t = \sum_i \operatorname{diam}(U_i)^s \operatorname{diam}(U_i)^{t-s} \le \delta^{t-s} \sum_i \operatorname{diam}(U_i)^s.$$

Taking the infimum over all countable δ-covers of A yields

$$\mathcal{H}^t_\delta(A) \le \delta^{t-s}\mathcal{H}^s_\delta(A).$$

If $\mathcal{H}^s(A) < \infty$, we find that

$$\mathcal{H}^t(A) = \lim_{\delta\to 0^+} \mathcal{H}^t_\delta(A) \le \lim_{\delta\to 0^+} \delta^{t-s}\mathcal{H}^s_\delta(A) = \lim_{\delta\to 0^+} \delta^{t-s}\mathcal{H}^s(A) = 0.$$

The second implication is the contraposition of the first. □

Definition 4.5. *The* **Hausdorff dimension** *of* X *is defined as*

$$\dim_{\mathrm{H}}(X) := \inf\{s \ge 0\colon \mathcal{H}^s(X) = 0\} = \sup\{s \ge 0\colon \mathcal{H}^s(X) = \infty\} \in [0,\infty]$$

with the usual conventions $\inf(\emptyset) = \infty$ *and* $\sup(\emptyset) = 0$.

Remark 4.1. *If* $s := \dim_{\mathrm{H}}(X)$ *denotes the Hausdorff dimension of* X, *then the* s*-dimensional Hausdorff measure of* X *could be equal to* 0, ∞ *or any value in between. On the other hand, if we can show that* $0 < \mathcal{H}^s(X) < \infty$ *for some* s, *then by Proposition* 4.2 *we know that* $s = \dim_{\mathrm{H}}(X)$.

Example 4.3 (Example 2.7 in [7])**.** *Let us prove that the Hausdorff dimension of the Cantor set* $\mathcal{C}$ *is*

$$\dim_{\mathrm{H}}(\mathcal{C}) = \frac{\log(3)}{\log(2)}.$$

We do this by showing that

$$\frac{1}{2} \le \mathcal{H}^s(\mathcal{C}) \le 1$$

for $s := \frac{\log(3)}{\log(2)}$ *and using Remark* 4.1.

- $\frac{1}{2} \le \mathcal{H}^s(\mathcal{C})$:
  Let $\delta \ge 0$ *and* $\{U_i\}_i$ *be a countable* δ*-cover of* $\mathcal{C}$. *By replacing* U_i *by the interval* $[\inf(U_i), \sup(U_i)]$ *we do not increase the diameter and hence end up with a countable* δ*-cover of* $\mathcal{C}$ *which is made up of closed intervals. So*

we may assume w.l.o.g. that $\{U_i\}_i$ *is a countable* δ*-cover of closed interval. Furthermore, as every interval* U_i *satisfies* $\operatorname{diam}(U_i) < \delta$ *we can slightly expand the intervals to get open intervals* $V_i \supseteq U_i$ *which, for some small* $\varepsilon > 0$*, satisfy*

$$\operatorname{diam}(V_i) < \min\{\delta, (1+\varepsilon)\operatorname{diam}(U_i)\}.$$

Since $\mathcal{C}$ *is compact, there exists a finite subcover, say* $\{V_1, \ldots, V_n\}$*. Taking the closure of* $V_1, \ldots, V_n$ *yields a finite* δ*-cover of closed interval for which*

$$\sum_{i=1}^{n} \operatorname{diam}(V_i)^s \leq (1+\varepsilon)^s \sum_i \operatorname{diam}(U_i)^s.$$

We may therefore assume w.l.o.g. that $\{U_i\}_i$ *is a finite* δ*-cover of closed intervals.*

So w.l.o.g. let $\{U_1, \ldots, U_n\}$ *be a finite* δ*-cover of* $\mathcal{C}$ *consisting of closed intervals. For* $1 \leq i \leq n$ *let* $k_i \in \mathbb{N}$ *such that*

$$3^{-(k_i+1)} \leq \operatorname{diam}(U_i) < 3^{-k_i}.$$

As all intervals in C_{k_i} *are at least a distance of* 3^{-k_i} *apart from each other,* U_i *intersects at most one of the intervals in* C_{k_i}*. Set* $j := \max\{k_1, \ldots, k_n\}$ *so that*

$$\forall 1 \leq i \leq n \colon 3^{-(j+1)} \leq \operatorname{diam}(U_i).$$

Form each interval in C_{k_i} *we get* 2^{j-k_i} *intervals in* C_j*. Therefore,* U_i *intersects at most*

$$2^{j-k_i} = 2^j 3^{-sk_i} = 2^j 3^s \left(3^{-(k_i+1)}\right)^s \leq 2^j 3^s \operatorname{diam}(U_i)^s$$

intervals in C_j*. Since there are* 2^j *intervals in* C_j *and each interval must be intersected by some* U_i *we find that*

$$2^j \leq \sum_{i=1}^{n} 2^{j-k_i} \leq \sum_{i=1}^{n} 2^j 3^s \operatorname{diam}(U_i).$$

Multiplying by $2^{-j}3^{-s}$ *yields*

$$3^{-s} \leq \sum_{i=1}^{n} \operatorname{diam}(U_i)^s.$$

Since this holds for all δ*-coverings, we find* $3^{-s} \leq \mathcal{H}^s_\delta(C)$ *and hence also*

$$\mathcal{H}^s(C) = \lim_{\delta \to 0^+} \mathcal{H}^s_\delta(C) \geq 3^{-s} = \frac{1}{2}.$$

- $\mathcal{H}^s(C) \leq 1$*:*
  Using the construction of the Cantor set, it is clear that for all $n \in \mathbb{N}$*, the* 2^n *Intervals which make up* C_n *form a countable* (3^{-n})*-cover of* $\mathcal{C}$*. Therefore,*

$$\mathcal{H}^s_{3^{-n}}(C) \leq \sum_{k=1}^{2^n} (3^{-n})^s = 2^n 3^{-ns} = 1$$

  and thus

$$\mathcal{H}^s(C) = \lim_{n \to \infty} \mathcal{H}^s_{3^{-n}}(C) \leq 1.$$

As seen in Example 4.3, computing the Hausdorff dimension by using the trick in Remark 4.1 can be very time consuming. In particular, proving the lower bound $0 < \mathcal{H}^s(X)$ is often a bit tedious. However, there is an easier way to compute the Hausdorff dimension of self-similar fractals which do not 'overlap too much in their fractal construction'. We need to introduce some more terminology to make this notion precise.

Definition 4.6. *A map* $f\colon X \to X$ *is called a* **contraction** *of* X *if*

$$\forall x, y \in X\colon d(f(x), f(y)) \leq c d(x, y),$$

for some $0 < c < 1$*. If we have equality*

$$\forall x, y \in X\colon d(f(x), f(y)) = c d(x, y),$$

we call f *a* **contracting similarity** *and* c *the* **ratio** *of* f*.*

Definition 4.7 (Open set condition)**.** *A family of contractions* $\{f_1, \ldots, f_m\}$ *on* X *satisfies the* **open set condition** *if there exists a non-empty open set* $V \subseteq X$ *that satisfies*

$$V \supseteq \bigsqcup_{i=1}^{m} f_i(V),$$

where $f_i(V)$ *are disjoint.*

The following theorem makes computing the Hausdorff dimension of self-similar fractals much easier, and shows that in many cases the Minkowski and Hausdorff dimensions agree.

Theorem 4.1 (Theorem 9.3 in [7])**.** *Let* $\{f_1, \ldots, f_m\}$ *be contracting similarities on a complete metric space with ratios* $0 < c_1, \ldots, c_m < 1$ *which satisfy the open set condition and consider the attractor*

$$F = \bigcup_{i=1}^{m} f_i(F).$$

Then $\dim_{\mathrm{H}}(F) = \dim_{\mathrm{mink}}(F) = s$*, where* s *is given by*

$$\sum_{i=1}^{m} c_i^s = 1.$$

Moreover,

$$0 < \mathcal{H}^s(F) < \infty.$$

We refer to [7] for a proof of Theorem 4.1.

Example 4.4 (Example 4.3 revisited)**.** *Let us verify Theorem* 4.1 *on the Cantor set* $\mathcal{C}$*. If we consider* $\mathcal{C}$ *on* $[0, 1]$*, the maps* $f_1, f_2\colon [0, 1] \to [0, 1]$ *given by*

$$f_1(x) := \frac{1}{3}x,\ f_2(x) := \frac{1}{3}x + \frac{2}{3}$$

are contracting similarities with ratios $c_1, c_2 = \frac{1}{3}$*. They satisfy the open set condition for* $V := (0, 1)$ *as*

$$(0, 1) \supseteq \left(0, \frac{1}{3}\right) \sqcup \left(\frac{2}{3}, 1\right) = f_1((0, 1)) \sqcup f_2((0, 1)).$$

The Cantor set is the attractor

$$\mathcal{C} = f_1(\mathcal{C}) \cup f_2(\mathcal{C}).$$

So by Theorem 4.1*, the Hausdorff dimension* $s \coloneqq \dim_{\mathrm{H}}(\mathcal{C})$ *is given by*

$$\left(\frac{1}{3}\right)^s + \left(\frac{1}{3}\right)^s = 1.$$

Solving for s *confirms our result from Example* 4.3*,*

$$s = \dim_{\mathrm{H}}(\mathcal{C}) = \frac{\log(3)}{\log(2)}.$$

4.3 Minimal Spanning Tree Dimension

In this section we define the MST dimension and show that it is equal to the Minkowski dimension following [9] closely. This is in preparation for Chapter 5, where we define other fractal dimensions that build on the idea of MSTs and the MST dimension.
Let $A \subseteq X$ be a finite subspace. For a MST $T(A)$ on A and for $\alpha \geq 0$ consider the quantity

$$E_\alpha(A) \coloneqq \sum_{e \in T(A)} \|e\|^\alpha,$$

where $\|e\| = d(x, y)$ denotes the length of the edge $e \in T(A)$ between $x, y \in A$.

Definition 4.8. *The* **minimal spanning tree dimension (MST dimension)** *of* X *is defined as*

$$\dim_{\mathrm{MST}}(X) \coloneqq \inf\{\alpha \geq 0 \colon \exists C \in \mathbb{R} \colon \forall A \subseteq X \textit{ finite} \colon E_\alpha(A) < C\}.$$

Theorem 4.2 (Theorem 2 in [9])**.** *We have*

$$\dim_{\mathrm{MST}}(X) = \dim_{\mathrm{mink}}(X).$$

Proof.

- $\dim_{\mathrm{mink}}(X) \leq \dim_{\mathrm{MST}}(X)$:
If $\dim_{\mathrm{mink}}(X) = 0$, then we are done. So suppose $\dim_{\mathrm{mink}}(X) > \alpha > 0$. Then

$$\begin{aligned}\log\left(\limsup_{\varepsilon \to 0^+} N(X,\varepsilon)\varepsilon^\alpha\right) &= \limsup_{\varepsilon \to 0^+} \left(\log(N(X,\varepsilon)) + \alpha \log(\varepsilon)\right) \\ &= \limsup_{\varepsilon \to 0^+} \log\left(\frac{1}{\varepsilon}\right)\left(\frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} + \alpha \frac{\log(\varepsilon)}{\log\left(\frac{1}{\varepsilon}\right)}\right) \\ &= \limsup_{\varepsilon \to 0^+} \log\left(\frac{1}{\varepsilon}\right)\left(\frac{\log(N(X,\varepsilon))}{\log\left(\frac{1}{\varepsilon}\right)} - \alpha\right) = \infty.\end{aligned}$$

So there exists a sequence $\{\varepsilon_n\}_n$ such that

$$\varepsilon_n \to 0^+, \ N(X,\varepsilon_n)\varepsilon_n^\alpha \to \infty \quad \text{as } n \to \infty.$$

Denote by $\{B(x_i^n, \varepsilon_n)\}_i$ a maximal ε_n-packing of X and consider the finite subset $A^n := \{x_i^n\}_i \subseteq X$. Every edge e in a MST on A^n has length $\|e\| \geq 2\varepsilon_n$ as $\{B(x_i^n, \varepsilon_n)\}_i$ are disjoint. Therefore

$$E_\alpha(A^n) = \sum_{e \in T(A^n)} \|e\|^\alpha \geq \sum_{e \in T(A)} 2^\alpha \varepsilon_n^\alpha = 2^\alpha (N(X, \varepsilon_n) - 1)\varepsilon_n^\alpha.$$

Since $N(X, \varepsilon_n)\varepsilon_n^\alpha \to \infty$, the sequence $\{E_\alpha(A^n)\}_n$ is unbounded and hence $\dim_{\text{MST}}(X) > \alpha$ which implies

$$\dim_{\text{mink}}(X) \leq \dim_{\text{MST}}(X).$$

- $\dim_{\text{MST}}(X) \leq \dim_{\text{mink}}(X)$:
  Let $\alpha > \dim_{\text{mink}}(X)$. Then there exists $C \in \mathbb{R}$ such that

$$\forall \varepsilon > 0\colon N(X, \varepsilon) \leq C\varepsilon^{-\alpha}.$$

Let $A \subseteq X$ be finite and consider a MST constructed by the following greedy algorithm: Start the tree with any point and iteratively add a point which is closest to the tree and connect it to the point which it is closest to. Now construct a $\left(\frac{\varepsilon}{2}\right)$-packing of X as follows: For every edge $e \in T(A)$ with $\|e\| \geq \varepsilon$, add the ball $B(x, \frac{\varepsilon}{2})$ where x is the point added to T when creating e. These balls form a $\left(\frac{\varepsilon}{2}\right)$-packing on X. Denote the diameter of X by $d := \operatorname{diam}(X)$. Then, for any $\beta > \alpha$,

$$\begin{aligned} E_\beta(A) &= \sum_{e \in T(A)} \|e\|^\beta \\ &= \sum_{n=0}^{\infty} \sum_{\substack{e \in T(A) \\ d2^{-(n+1)} < \|e\| \leq d2^{-n}}} \|e\|^\beta \\ &\leq \sum_{n=0}^{\infty} N(X, d2^{-(n+1)})(d2^{-n})^\beta \\ &\leq \sum_{n=0}^{\infty} C(d2^{-(n+1)})^{-\alpha}(d2^{-n})^\beta \\ &= Cd^{\beta-\alpha}2^\alpha \sum_{n=0}^{\infty} 2^{n(\alpha-\beta)}. \end{aligned}$$

Therefore, $E_\beta(A)$ is bounded, which means $\dim_{\text{MST}}(X) \leq \alpha$.

$\square$

Chapter 5

Persistent Homology Dimension

In this chapter we approximate finite metric spaces with Vietoris-Rips complexes and study their homology. This yields a persistence module that enables us to analyse the topological features of the space. For a detailed treatment of persistent homology, we refer the reader to [3]. By showing the correspondence between the lifetimes intervals in the 0th persistence module and edges of a MSTs (see Proposition 5.1) we generalise fractal dimensions based on MSTs. We do this using two different approaches. In Section 5.2.1 we follow [18] and study persistence modules on extremal subsets. In Section 5.2.2 we follow [2] and analyse persistence modules on random subsets.

5.1 Persistent Homology

Persistent homology is a tool from algebraic topology that can be used to capture topological features of finite metric spaces and to describe the behaviour of these features over different scales. This concept can be used to analyse data samples and draw conclusions about the structure of the underlying space.
For the entirety of this chapter, let (X, d) denote a compact metric space. In order to analyse the topological features of X, we construct an abstract simplicial complex on X and study its homology.

Definition 5.1. *An* **abstract simplicial complex** *on X is a family $\mathcal{F}$ of non-empty finite subsets which is closed under taking non-empty subsets, that is,*

$$\forall A \in \mathcal{F} \colon \emptyset \neq B \subseteq A \implies B \in \mathcal{F}.$$

Definition 5.2. *The* **Vietoris-Rips complex** $\mathrm{VR}(X, r)$ *for a given* **scale parameter** $r \geq 0$ *is the abstract simplicial complex whose vertices are the points in X and whose k-simplices consist exactly of all $(k+1)$ point sets $\{x_0, \ldots, x_k\} \subseteq X$, which have diameter* $\mathrm{diam}(\{x_0, \ldots, x_k\}) \leq r$.

Example 5.1. *Let X be a metric space depicted in Figure 5.1 obtained by sampling* 20 *points from an annulus. Figure 5.2 shows the Vietoris-Rips complexes* $\mathrm{VR}(X, r)$ *for eight different scale parameters r.*

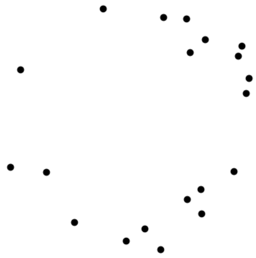


Figure 5.1: 20 points sampled from an annulus.

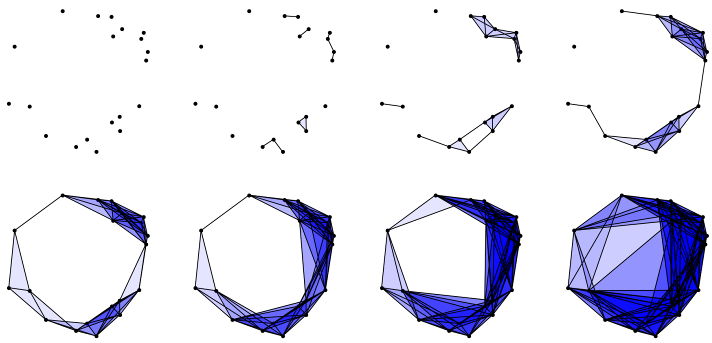


Figure 5.2: Example of Vietoris-Rips complexes.

For $0 \leq r < r'$ there is a natural inclusion

$$\mathrm{VR}(X, r) \subseteq \mathrm{VR}(X, r')$$

as every k-simplex in $\mathrm{VR}(X, r)$ is also a k-simplex in $\mathrm{VR}(X, r')$. This inclusion induces a linear map on the homology groups

$$H_k(\mathrm{VR}(X, r)) \to H_k(\mathrm{VR}(X, r')).$$

The homology groups may be taken over an arbitrary field $\mathbb{F}$, which we drop from notation.

Definition 5.3. *By letting the scale parameter r range over $\mathbb{R}_{\geq 0}$, we obtain the* **Vietoris–Rips filtration**

$$\{\mathrm{VR}(X, r)\}_{r \in \mathbb{R}_{\geq 0}}.$$

The family of homology groups

$$\{H_k(\mathrm{VR}(X, r))\}_{r \in \mathbb{R}_{\geq 0}}$$

together with the induced linear maps between them is called the **kth persistence module**.

We have the following decomposition theorem for the kth persistent module.

Theorem 5.1. *If X is finite, then there exist $N \in \mathbb{N}$, $\{b_0, \dots, b_N\} \in [0, \infty)$, $\{d_0, \dots, d_N\} \in [0, \infty]$ with $b_n \leq d_n$ and in isomorphism*

$$\{H_k(\mathrm{VR}(X, r))\}_{r \in \mathbb{R}_{\geq 0}} \cong \bigoplus_{n=0}^{N} U(b_n, d_n),$$

where

$$U(b_n, d_n)_r \coloneqq \begin{cases} \mathbb{F}, & \text{if } r \in [b_n, d_n], \\ 0, & \text{else.} \end{cases}$$

The proof uses the following classification theorem for finitely generated $\mathbb{F}[t]$-modules.

Theorem 5.2 (Theorem 2.10 in [3])**.** *Let M_* be a finitely generated $\mathbb{F}[t]$-module. Then there exist integers $\{i_1, \dots, i_m\}, \{j_1, \dots, j_n\}, \{l_1, \dots, l_n\}$, and an isomorphism*

$$M_* \cong \bigoplus_{r=1}^{m} \mathbb{F}[t](i_r) \oplus \bigoplus_{s=1}^{n} (\mathbb{F}[t]/(t^{l_s}))(j_s).$$

This decomposition is unique up to permutation of factors.

Proof of Theorem 5.1. If X is finite, then the Vietoris-Rips filtration yields only finitely many distinct complexes. Let $r_0 < \ldots < r_M$ denote the scale parameter of these different complexes. This allows us to view the $\mathbb{R}$-persistence module as an $\mathbb{N}$-persistence module

$$\{H_k(\mathrm{VR}(X, r))\}_{r \in \mathbb{R}_{\geq 0}} \cong \{H_k(\mathrm{VR}(X, r_n))\}_{n \in \mathbb{N}},$$

where $r_n \coloneqq r_M$ for all $n > M$.

Claim 5.1. *There is an equivalence between the category of $\mathbb{N}$-persistence abelian groups $\mathbb{N}_{\mathrm{pers}}(\underline{\mathsf{Ab}})$ and the category of non-negatively graded modules over $\mathbb{Z}[t]$.*

Proof of Claim. Let $\{A_n\}_{n \in \mathbb{N}}$ be an $\mathbb{N}$-persistence abelian group with maps

$$\psi_{m,n} \colon A_m \to A_n.$$

Its associated graded module over $\mathbb{Z}[t]$ is given by

$$\theta(\{A_n\}_{n \in \mathbb{N}}) \coloneqq \bigoplus_{n \in \mathbb{N}} A_n,$$

where multiplication with t on elementary elements $\alpha_n \in A_n$ is given by

$$t \cdot \alpha_n \coloneqq \psi_{n,n+1}(\alpha_n).$$

θ is a functor from $\mathbb{N}_{\mathrm{pers}}(\underline{\mathsf{Ab}})$ to the category of non-negatively graded modules over $\mathbb{Z}[t]$. Its inverse is given by defining

$$\psi_{m,n}(\alpha_m) \coloneqq t^{n-m} \alpha_m.$$

□

By applying the functor θ from Claim 5.1 to the $\mathbb{N}$-persistence module we get a $\mathbb{Z}[t]$-module

$$M_* := \theta\left(\{H_k(\mathrm{VR}(X, r_n))\}_{n\in\mathbb{N}}\right).$$

In fact, since the homology groups are taken over a field $\mathbb{F}$, M_* is also a $\mathbb{F}[t]$-module. Since $H_k(\mathrm{VR}(X, r_n))$ is finite dimensional for all $n \in \mathbb{N}$ and $\mathrm{VR}(X, r_n) = \mathrm{VR}(X, r_M)$ for all $n > M$, M_* is finitely generated. This means we can apply the classification theorem 5.2 for finitely generated $\mathbb{F}[t]$-modules

$$M_* \cong \bigoplus_{r=1}^{m} \mathbb{F}[t](i_r) \oplus \bigoplus_{s=1}^{n} (\mathbb{F}[t]/(t^{l_s}))(j_s).$$

Using the following correspondence and switching back to indexing over $\mathbb{R}$ concludes the proof.

$$\mathbb{F}[t](i_r) \cong U(i_r, \infty),$$
$$(\mathbb{F}[t]/(t^{l_s}))(j_s) \cong U(i_r, i_r + l_s).$$

□

Definition 5.4. *Suppose X is finite and decompose the kth persistence module according to Theorem* 5.1

$$\{H_k(\mathrm{VR}(X, r))\}_{r\in\mathbb{R}_{\geq 0}} \cong \bigoplus_{n=0}^{N} U(b_n, d_n).$$

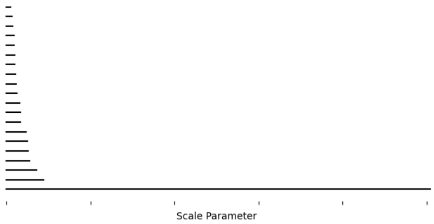
We call b_n and d_n the **birth** *and* **death scale** *of a feature, and $[b_n, d_n]$ the* **lifetime interval**. *The multiset of lifetime intervals is denoted by* $\mathrm{PH}_k(X)$.

This representation allows us to visualise the persistence module in the form of a ***barcode*** or a ***persistence diagram***. The barcode of a metric space is obtained by plotting the lifetime intervals in $\mathrm{PH}_k(X)$ over $[0, \infty]$. The persistence diagram is obtained by plotting the tuples of life and death scales of each generator in $[0, \infty) \times (0, \infty]$.

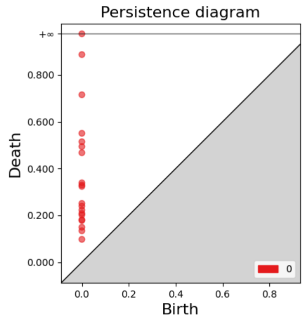
Example 5.2. *Consider again the metric space X from Example* 5.1*,* 20 *points sampled from a annulus. Figure 5.3 shows the barcode and persistence diagram for* $\mathrm{PH}_0(X)$*, while Figure 5.4 shows the barcode and persistence diagram for* $\mathrm{PH}_1(X)$*. We use [14, 5] for the computation and plotting.*

We can analyse the persistent homology of a sample to deduces the topological features of the underlying space. Longer lifetime intervals correspond to a generator which has persisted over a larger range of scale and are hence likely to indicate an actual topological feature of the underlying space. Conversely, shorter lifetime intervals indicate that the corresponding generator has not persisted very long and has likely only appeared as an artefact from sampling.

Example 5.3 (Continuing Example 5.2)**.** *As we can see in the barcode and the persistence diagram of* $\mathrm{PH}_1(X)$ *in Figure 5.4, there are two lifetime intervals. One is very short whereas the other is significantly longer. When looking at the Vietoris-Rips complexes in Figure 5.2, we see that the generator corresponding to the shorter lifetime interval is only visible in the third complex. The generator has appeared due to the sampling and does not describe the topology of the*



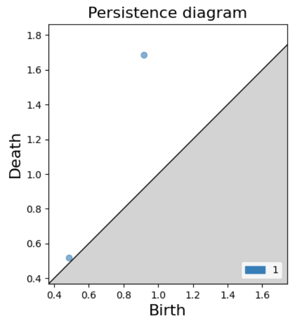
(a) Barcode



(b) Persistence diagram

Figure 5.3: Barcode and persistence diagram for PH_0.

(a) Barcode



(b) Persistence diagram

Figure 5.4: Barcode and persistence diagram for PH_1.

annulus. On the other hand, the generator corresponding to the longer lifetime interval is present in the fifth, sixth and seventh Vietoris-Rips complex. It appeared due to the hole in the annulus and is describing a topological feature of the under space.

5.2 Persistent Homology Dimension

In this section we establish the relation between $\mathrm{PH}_0(X)$ a MST X. We use this to generalise fractal dimension based on MSTs.

Proposition 5.1. *There is a correspondence of bounded lifetime intervals in* $\mathrm{PH}_0(X)$ *and edges in a MST on* X*, where the length of a lifetime interval is equal to the length of its corresponding edge.*

Proof. Let us see how the 0th homology of the Vietoris-Rips complex changes as we increase the scale parameter. It is enough to only consider 0 and 1-simplices of $\mathrm{VR}(X, r)$. Note that the connected components of $\mathrm{VR}(X, r)$ are generators of $H_0(\mathrm{VR}(X, r))$ and that by increasing the scale parameter r, the number of connected components in $\mathrm{VR}(X, r)$ cannot increase, so no new generators can appear. For $r = 0$, the Vietoris-Rips complex $\mathrm{VR}(X, 0)$ consists of only the vertices in X. Each point in X is a generator of $H_0(\mathrm{VR}(X, 0))$ and has birth scale 0. By increasing r, we start adding 1-simplices to the Vietoris-Rips complex between two points x and y as soon as $d(x, y) = r$. Whenever such a 1-simplex connects two previously disconnected components, the number of connected components decreases by one and a generator disappears with death scale is equal to r. Its lifetime interval is $[0, r]$ and has length r. The 1-simplices which connects two previously disconnected components correspond exactly to the edges one obtains by Kruskal's algorithm (see Section 2.2), proving the statement. □

This correspondence can be used to generalise fractal dimensions that are defined via MSTs. There have been multiple different approaches to this, two of which we are going to analyse here.

- In Section 5.2.1, we follow the approach taken in [18]. We define a persistent homology dimension for metric spaces which generalises a fractal dimension that is based on MSTs on extremal subsets. This definition has the advantage of being equal to the Minkowski dimension in case of the 0th persistent homology (see Theorem 5.3). The disadvantage of this definition is that it is difficult to compute.
- In Section 5.2.2, we follow the approach taken in [2]. We define a persistent homology dimension for probability measures on metric spaces which generalises a fractal dimension based on MSTs on random subsets. The advantage of this definition is that it is easy to compute an estimate. As of now, no relations of to other fractal dimensions has been proven.

5.2.1 Persistent Homology Dimension (via Extremal Subsets)

The correspondence between the edges of a MST $T(A)$ on a finite subset $A \subseteq X$ and the bounded lifetime intervals in $\mathrm{PH}_0(A)$ from Proposition 5.1 lets us rewrite $E_\alpha(A)$ as

$$E_\alpha(A) = \sum_{(b,d) \in \mathrm{PH}_0(A)} (d - b)^\alpha,$$

where the sum is taken over all bounded lifetime intervals in $\mathrm{PH}_0(A)$. We can generalise this by looking at lifetime intervals in higher dimensional persistent homology.

Definition 5.5. *For $\alpha \geq 0$ and $k \in \mathbb{N}$ define*

$$E^k_\alpha(A) = \sum_{(b,d) \in \mathrm{PH}_k(A)} (d - b)^\alpha,$$

where the sum is taken over all bounded lifetime intervals in $\mathrm{PH}_k(A)$.

Using these newly defined quantities, we can also generalise the MST dimension from Definition 4.8.

Definition 5.6. *For $k \in \mathbb{N}$, the* **k-dimensional persistent homology dimension** *is defined as*

$$\dim^k_{\mathrm{PH}}(X) \coloneqq \inf\{\alpha \geq 0 \colon \exists C \in \mathbb{R} \colon \forall A \subseteq X \text{ finite} \colon E^k_\alpha(A) < C\}.$$

Theorem 5.3. *We have*

$$\dim^0_{\mathrm{PH}}(X) = \dim_{\mathrm{mink}}(X).$$

Proof. By the correspondence of bounded intervals in $\mathrm{PH}_0(X)$ and edges in a MST on X described in Proposition 5.1, we have $E_\alpha(A) = E^0_\alpha(A)$ and thus

$$\dim^0_{\mathrm{PH}}(X) = \dim_{\mathrm{MST}}(X).$$

With Theorem 4.2, we find that

$$\dim^0_{\mathrm{PH}}(X) = \dim_{\mathrm{mink}}(X).$$

□

5.2.2 Persistent Homology Dimension (via Random Subsets)

Consider a probability measure μ on X and let X_n be a of n points from X according to μ. Denote by $T(X_n)$ a MST on X_n and consider the total length of the MST

$$L(X_n) \coloneqq \sum_{e \in T(X_n)} \|e\|.$$

Steele has analysed the asymptotic behaviour of $L(X_n)$ in [19]. By the correspondence between bounded intervals in $\mathrm{PH}_0(X_n)$ and edges in $T(X_n)$ described in Proposition 5.1 we can rewrite the total length of $T(X_n)$ as

$$L(X_n) = \sum_{(b,d) \in \mathrm{PH}_0(X_n)} (d-b),$$

where the sum is taken over all bounded lifetime intervals in $\mathrm{PH}_0(X_n)$. We can generalise this quantity by looking at lifetime intervals in higher dimensional persistent homology.

Definition 5.7. *For $k \in \mathbb{N}$ define*

$$L_k(X_n) \coloneqq \sum_{(b,d) \in \mathrm{PH}_k(X_n)} (d-b),$$

where the sum is taken over all bounded lifetime intervals in $\mathrm{PH}_k(X_n)$.

We can now analyse the asymptotic behaviour of $L_k(X_n)$ as $n \to \infty$ and define another fractal dimension.

Definition 5.8. *Denote by $X_n \subseteq X$ an i.i.d. sample of n points from X according to a probability measure μ. The* **k-dimensional persistent homology dimension** *of μ is defined as*

$$\dim^k_{\mathrm{PH}}(\mu) \coloneqq \inf\left\{d > 0 \colon \exists C \in \mathbb{R} \colon \lim_{n\to\infty} \mathbb{P}\left[L_k(X_n) \leq C n^{\frac{d-1}{d}}\right] = 1\right\}.$$

This definition allows us to approximate the k-dimensional persistent homology fractal dimension:
Write $d := \dim_{\mathrm{PH}}^k(\mu)$ and $\alpha := \frac{d-1}{d}$. Almost surely, the asymptotic growth of $L_k(X_n)$ is

$$L_k(X_n) \sim Cn^{\alpha}, \quad \text{as } n \to \infty.$$

Taking the log yields

$$\log(L_k(X_n)) \sim \log(C) + \alpha \log(n), \quad \text{as } n \to \infty.$$

To estimate d, we take a increasing sequence of sample sizes $\{n_i\}_i$, and compute the quantities $L_k(X_{n_i})$. We plot these vales on a log-log scale and look at the asymptotic part, which looks like a straight line with slope α. We estimate α and get an estimate for d by

$$d = \frac{1}{1-\alpha}.$$

Example 5.4. *Let us estimate the persistent homology dimension for the uniform distribution μ on the Sierpinski triangle $\mathcal{S}$. As a first approximation, let us consider S_5 from the construction of the Sierpinski triangle (see Example* 3.2*) instead of the Sierpinski triangle itself. We take uniformly distributed samples of up to* 10000 *points from S_5. Figure 5.5 show the log-log plot of L_0 for these samples. The asymptotic slope is approximately*

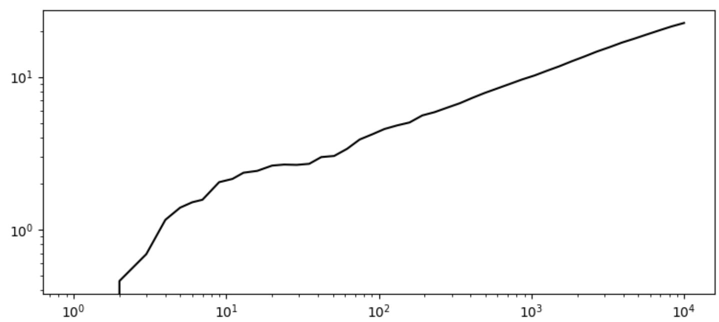


Figure 5.5: $L_0(X_n)$ for samples from the Sierpinski triangle.

$$\alpha \approx 0.3187.$$

Therefore we can an approximate persistent homology dimension as

$$\dim_{\mathrm{PH}}(\mu) \approx \frac{1}{1-\alpha} \approx 1.4679.$$

This is close to the Minkowski and Hausdorff dimension of the Sierpinski triangle which is

$$\dim_{\mathrm{mink}}(\mathcal{S}) = \dim_{\mathrm{H}}(\mathcal{S}) = \frac{\log(3)}{\log(2)} \approx 1.5850.$$

Chapter 6

Magnitude Dimension

In this chapter we present the notion of magnitude. We first define magnitude for matrices and finite metric spaces following [11]. Then we define magnitude for infinite metric spaces referring to [20], [15] and [16]. In Section 6.2 we give a continuity statement about magnitude which enables us to approximate magnitude of infinite metric space. This is based on [12]. Finally, we define the magnitude function which tells us how magnitude of a metric space changes as we scale the space. By analysing the growth rate of the magnitude function, we define another fractal dimension – the magnitude dimension. We show that for compact subspace of $\mathbb{R}^n$, the magnitude and Minkowski dimension agree (see Theorem 6.4), referring to [16].

6.1 Magnitude

The aim of this section is to introduce magnitude. We first define magnitude for matrices before using the associated similarity matrix to define magnitude for finite metric spaces. For this part, we follow [11] closely. In Section 6.1.4 we extend this definition to infinite metric spaces, mainly following [20].

6.1.1 Magnitude of Matrices

Let k be a commutative semiring with unit $1 \in k$ and X be a non-empty, finite set. Let $Z \in k^{X \times X}$ be a square matrix over k indexed by elements in X.

Definition 6.1. *A column vector $w \in k^X$ is called a* **weighting** *of Z if $Zw = \mathbb{1}$, where $\mathbb{1} \in k^X$ denote the column vector with all 1s. A row vector $v \in k^X$ is called a* **coweighting** *of Z if $vZ = \mathbb{1}^{\mathrm{T}}$.*

Remark 6.1. *It is possible for a matrix to admit a weighting but not coweighting or vice versa. For example*

$$Z := \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

has a weighting

$$w := \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

but Z does not admit a coweighting.

Definition 6.2. *If $Z \in k^{X\times X}$ admits both a weighting $w \in k^X$ and a coweighting $v \in k^X$, then*

$$\sum_{x\in X} w(x) = \mathbb{1}^{\mathrm{T}} w = vZw = v\mathbb{1} = \sum_{x\in X} v(x). \tag{6.1}$$

In this case, we define the **magnitude** *$|Z|$ of Z to be this sum*

$$|Z| \coloneqq \sum_{x\in X} w(x) = \sum_{x\in X} v(x).$$

By (6.1) the magnitude of a matrix is independent of the weighting and coweighting. Hence, the magnitude of a matrix is well defined.

Proposition 6.1 (Lemma 1.1.4 in [11])**.** *If Z is invertible, then there exists a unique weighting and a unique coweighting. Hence, Z admits a magnitude which is given by the sum of all entries of Z^{-1},*

$$|Z| = \sum_{x,y\in X} Z^{-1}(x,y).$$

Proof. The unique weighting and coweighting are given by $w \coloneqq Z^{-1}\mathbb{1}$ and $v \coloneqq \mathbb{1}^{\mathrm{T}} Z^{-1}$ respectively. Therefore Z admits a magnitude which is given by

$$|Z| = \sum_{x\in X} w(x) = \sum_{x\in X} (Z^{-1}\mathbb{1})(x) = \sum_{x,y\in X} Z^{-1}(x,y).$$

□

Proposition 6.2 (Proposition 2.4.3 in [11])**.** *If $Z \in \mathbb{R}^{X\times X}$ is positive definite, then Z admits a magnitude which is given by*

$$|Z| = \sup_{u\in\mathbb{R}^X\setminus\{0\}} \frac{\left(\sum_{x\in X} u(x)\right)^2}{u^{\mathrm{T}} Z u} = \max_{u\in\mathbb{R}^X\setminus\{0\}} \frac{\left(\sum_{x\in X} u(x)\right)^2}{u^{\mathrm{T}} Z u}.$$

The supremum is attained exactly when $u \in \mathbb{R}^X \setminus \{0\}$ is a scalar multiple of the unique weighting of Z.

Proof. Since Z is positive definite, Z is invertible. By Proposition 6.1 Z has a unique weighting $w \coloneqq Z^{-1}\mathbb{1}$ and admits a magnitude.
We have

$$\frac{\left(\sum_{x\in X} w(x)\right)^2}{w^{\mathrm{T}} Z w} = \frac{\left(\mathbb{1}^{\mathrm{T}} w\right)^2}{\mathbb{1}^{\mathrm{T}} w} = \mathbb{1}^{\mathrm{T}} w = |Z|.$$

Let us now show that for all $u \in \mathbb{R}^X \setminus \{0\}$

$$\frac{\left(\sum_{x\in X} u(x)\right)^2}{u^{\mathrm{T}} Z u} \leq \frac{\left(\sum_{x\in X} w(x)\right)^2}{w^{\mathrm{T}} Z w}.$$

Since Z is positive definite,

$$\mathbb{R}^X \times \mathbb{R}^X \to \mathbb{R},\ (u, \tilde{u}) \mapsto u^{\mathrm{T}} Z \tilde{u}$$

defines an inner product on $\mathbb{R}^X$. We can rewrite

$$\begin{aligned}
\frac{\left(\sum_{x\in X} u(x)\right)^2}{u^{\mathrm{T}} Z u} &\leq \frac{\left(\sum_{x\in X} w(x)\right)^2}{w^{\mathrm{T}} Z w} \\
\left(w^{\mathrm{T}} Z w\right)\left(\mathbb{1}^{\mathrm{T}} u\right)^2 &\leq \left(u^{\mathrm{T}} Z u\right)\left(\mathbb{1}^{\mathrm{T}} w\right)^2 \\
\left(w^{\mathrm{T}} Z w\right)\left(w^{\mathrm{T}} Z u\right)^2 &\leq \left(u^{\mathrm{T}} Z u\right)\left(w^{\mathrm{T}} Z w\right)^2 \\
\left(w^{\mathrm{T}} Z u\right)^2 &\leq \left(u^{\mathrm{T}} Z u\right)\left(w^{\mathrm{T}} Z w\right).
\end{aligned}$$

By the Cauchy-Schwarz inequality, this holds, and we have equality if and only if u is a non-zero scalar multiple of w. □

6.1.2 Magnitude of Finite Metric Spaces

Let (X, d) denote a finite metric space for the remainder of this section.

Definition 6.3. *The* **associated similarity matrix (ASM)** $Z_X \in \mathbb{R}^{X\times X}$ *of* X *is defined by*

$$Z_X(x, y) \coloneqq e^{-d(x,y)}.$$

A weighting or coweighting of Z_X, is sometimes simply referred to as a **weighting** *or* **coweighting** *of X respectively.*

Definition 6.4. *If Z_X admits a magnitude, that is, if X admits a weighting and a coweighting, then we define the* **magnitude** $|X|$ *of (X, d) to be the magnitude $|Z_X|$ of Z_X.*

Remark 6.2. *By the symmetry of the metric d, the ASM Z_X is a symmetric matrix. So if there exists a weighting $w \in \mathbb{R}^X$, then w^{T} is a coweighting as*

$$w^{\mathrm{T}} Z_X = (Z_X^{\mathrm{T}} w)^{\mathrm{T}} = (Z_X w)^{\mathrm{T}} = \mathbb{1}^{\mathrm{T}}.$$

Similarly, if $v \in \mathbb{R}^X$ is a coweighting, then v^T is a weighting.
Hence, admitting a weighting is a necessary and sufficient condition for finite metric spaces to admit a magnitude.

Example 6.1. *Let X be the metric space consisting of exacts two points at distance $d > 0$. Then the ASM is given by*

$$Z_X = \begin{pmatrix} 1 & e^{-d} \\ e^{-d} & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

This matrix is invertible with inverse

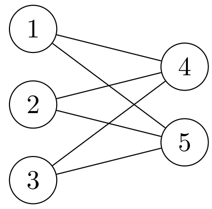
$$Z_X^{-1} = \begin{pmatrix} \frac{e^{2d}}{e^{2d}-1} & -\frac{e^{d}}{e^{2d}-1} \\ -\frac{e^{d}}{e^{2d}-1} & \frac{e^{2d}}{e^{2d}-1} \end{pmatrix}.$$

By Proposition 6.1*, the magnitude of X is given by*

$$|X| = 2\frac{e^{2d}}{e^{2d}-1} - 2\frac{e^{d}}{e^{2d}-1} = 1 + \tanh\left(\frac{d}{2}\right).$$

There are metric spaces, whose ASM do not admit a weighting, and hence, the metric space does not admit a magnitude.

Example 6.2 (Example 2.2.7 in [11])**.** *Consider the complete bipartite graph* $X \coloneqq K_{3,2}$ *with shortest path metric scaled by* $\log(\sqrt{2})$*.*



Its ASM is given by

$$Z_X = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{1}{2} & 1 & \frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{1}{2} & \frac{1}{2} & 1 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 1 & \frac{1}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2} & 1 \end{pmatrix} \in \mathbb{R}^{5\times 5}.$$

A quick computation shows that Z_X *and the augmented matrix* $[Z_X|\mathbb{1}]$ *have different rank. Therefore* X *does not admit a weighting, and hence* X *does not admit a magnitude.*

6.1.3 Magnitude of Compact, Positive Definite Metric Spaces

In this section we extend the definition of magnitude to infinite metric spaces via finite approximations.

Definition 6.5. *A finite metric space* (X, d) *is called* **positive definite** *if its ASM* Z_X *is positive definite. An infinite metric space* (X, d) *is called* **positive definite** *if every finite subspace is positive definite. We abbreviate 'positive definite metric space' by PDMS.*

Remark 6.3. *Since a principal submatrix of a positive definite matrix is also positive definite, every subspace of a PDMS is also positive definite.*

Proposition 6.3. *If* (X, d) *is a finite PDMS, then* X *admits a magnitude given by*

$$|X| = \sup_{u\in\mathbb{R}^X\setminus\{0\}} \frac{\left(\sum_{x\in X} u(x)\right)^2}{u^{\mathrm{T}} Z_X u} = \max_{u\in\mathbb{R}^X\setminus\{0\}} \frac{\left(\sum_{x\in X} u(x)\right)^2}{u^{\mathrm{T}} Z_X u}.$$

The supremum is attained exactly when $u \in \mathbb{R}^X \setminus \{0\}$ *is a scalar multiple of the unique weighting of* X*.*

Proof. This is an immediate consequence of Proposition 6.2. □

Definition 6.6. *We define the* **magnitude** $|X|$ *of a compact PDMS* (X, d) *to be*

$$|X| \coloneqq \sup\{|X'| \colon X' \subseteq X \text{ finite}\} \in [0, \infty].$$

Remark 6.4. *Since* X *is positively definite, every finite subspace* $X' \subseteq X$ *is so too, by definition. Hence, by Proposition* 6.3*,* X' *admits a magnitude, so* $|X|$ *is well-defined. As a consequence of Remark* 6.3 *and the following Proposition* 6.4*, Definitions* 6.4 *and* 6.6 *coincide on finite PDMS. So there is no conflict in terminology.*

Proposition 6.4 (Corollary 2.4.4 in [11]). *Suppose (X, d) is a finite PDMS and $X' \subseteq X$, then $|X'| \leq |X|$.*

Proof. For $u' \in \mathbb{R}^{X'} \setminus \{0\}$ define $u \in \mathbb{R}^X$ by

$$u(x) := \begin{cases} u'(x), & \text{if } x \in X', \\ 0, & \text{else.} \end{cases}$$

Since $u' \neq 0$, also $u \neq 0$ and furthermore

$$\frac{\left(\sum_{x\in X} u(x)\right)^2}{u^{\mathrm{T}} Z_X u} = \frac{\left(\sum_{x\in X'} u'(x)\right)^2}{u'^{\mathrm{T}} Z_{X'} u'}$$

and so the claim follows from Proposition 6.3. □

6.1.4 Magnitude via Measures

In this section we choose another approach to extend the definition of magnitude to infinite spaces. We introduce weight measures to extend the definition of weights following [20] and [16]. We refer the reader to [17] for a detailed treatment of measure theory.

Definition 6.7. *Denote by* $\mathrm{M}(X)$ *the* **space of finite signed Borel measures** *on X and by* $\mathrm{M}_+(X)$ *the* **space of finite non-negative Borel measures** *on X. The* **support** *of a non-negative measure $\mu \in \mathrm{M}_+(X)$ is given by*

$$\mathrm{supp}(\mu) := \{a \in X : \forall \varepsilon > 0 : \mu(B(a, \varepsilon)) > 0)\}.$$

Using the Hahn decomposition theorem, any signed measure $\mu \in \mathrm{M}(X)$ can be written as $\mu = \mu^+ - \mu^-$ with $\mu^+, \mu^- \in \mathrm{M}_+(X)$. The **support** *of μ is then defined as*

$$\mathrm{supp}(\mu) = \mathrm{supp}(\mu^+) \cup \mathrm{supp}(\mu^-).$$

Define by $\mathrm{FM}(X)$ *the* **space of finitely supported, finite signed measures** *on X.*

Definition 6.8. *We define a symmetric bilinear form $\langle \cdot, \cdot \rangle_{\mathrm{W}}$ on $\mathrm{FM}(X)$ by*

$$\langle \mu, \nu \rangle_{\mathrm{W}} := \int_X \int_X e^{-d(x,y)} \, \mathrm{d}\nu(y) \, \mathrm{d}\mu(x).$$

Lemma 6.1. *If (X, d) is a PDMS, then $\langle \cdot, \cdot \rangle_{\mathrm{W}}$ is an inner product on $\mathrm{FM}(X)$.*

Proof. Let $\mu \in \mathrm{FM}(X) \setminus \{0\}$. Denote by $B \subseteq X$ the finite support of μ. Since X is positive definite, the ASM of B is positive definite. Denote by $u \in \mathbb{R}^B \setminus \{0\}$ the vector given by $u(x) := \mu(\{x\})$. Therefore

$$\begin{aligned} \langle \mu, \mu \rangle &= \int_X \int_X e^{-d(x,y)} \, \mathrm{d}\mu(y) \, \mathrm{d}\mu(x) \\ &= \sum_{x\in B} \sum_{y\in B} e^{-d(x,y)} \mu(\{y\}) \mu(\{x\}) \\ &= u^{\mathrm{T}} Z_B u > 0. \end{aligned}$$

□

Definition 6.9. *Denote by* $\mathrm{W}(X)$ *the completion of* $\mathrm{FM}(X)$ *with respect to* $\langle\cdot,\cdot\rangle_{\mathrm{W}}$. *We call* $\mathrm{W}(X)$ *the* **weighting space** *of* X.

Suppose from now on that (X, d) is compact.

Definition 6.10. *A* **weight measure** *on* (X, d) *is a measure* $\mu \in \mathrm{W}(X)$ *such that for all* $x \in X$

$$\int_X e^{-d(x,y)}\,\mathrm{d}\mu(y) = 1.$$

Remark 6.5. *If* X *is finite, then* X *admits a weight measure* $\mu \in \mathrm{W}(X)$ *if and only if it admits a weighting* $w \in \mathbb{R}^X$ *as in Definition* 6.3*, where the correspondence is given by* $\mu(\{x\}) = w(x)$ *for all* $x \in X$.
Indeed, if X *is finite, for all* $x \in X$ *we have*

$$\int_X e^{-d(x,y)}\,\mathrm{d}\mu(y) = \sum_{y\in X} e^{-d(x,y)} w(y) = \sum_{y\in X} Z_X(x,y) w(y).$$

Definition 6.11. *If* X *admits a weight measure* $\mu \in \mathrm{W}(X)$*, we define the* **magnitude** $|X|$ *of* X *to be*

$$|X| \coloneqq \mu(X).$$

If X *does not admit a weight measure, we set* $|X| \coloneqq \infty$.

Remark 6.6. *The magnitude defined in Definition* 6.11 *is independent of the weight measure.*
Indeed, if $\mu, \nu \in \mathrm{W}(X)$ *are two weight measures on* X*, by Fubini we find*

$$\begin{aligned}
\mu(X) &= \int_X 1\,\mathrm{d}\mu(y) \\
&= \int_X \int_X e^{-d(x,y)}\,\mathrm{d}\nu(y)\,\mathrm{d}\mu(x) \\
&= \int_X \int_X e^{-d(x,y)}\,\mathrm{d}\mu(x)\,\mathrm{d}\nu(y) \\
&= \int_X 1\,\mathrm{d}\nu(x) \\
&= \nu(X).
\end{aligned}$$

Example 6.3 (Theorem 2 in [20])**.** *Let* $a, b \in \mathbb{R}$ *with* $a < b$ *and let us compute the magnitude of the compact interval* $[a, b] \subseteq \mathbb{R}$.

Claim 6.1. *A weight measure on* $[a, b]$ *is given by*

$$\mu \coloneqq \frac{1}{2}(\delta_a + \delta_b + \lambda),$$

where δ *and* λ *denote the Dirac and Lebesgue measure respectively.*

Proof of Claim. For all $x \in [a,b]$ we have

$$\begin{aligned}
2\int_{[a,b]} e^{-d(x,y)}\, \mathrm{d}\mu(y) &= \int_a^b e^{-d(x,y)}\, \mathrm{d}\delta_a(y) + \int_a^b e^{-d(x,y)}\, \mathrm{d}\delta_b(y) \\
&\quad + \int_a^b e^{-d(x,y)}\, \mathrm{d}\lambda(y) \\
&= e^{-d(x,a)} + e^{-d(x,b)} \\
&\quad + \int_a^x e^{-d(x,y)}\, \mathrm{d}\lambda(y) + \int_x^b e^{-d(x,y)}\, \mathrm{d}\lambda(y) \\
&= e^{a-x} + e^{x-b} + \int_a^x e^{y-x}\, \mathrm{d}\lambda(y) + \int_x^b e^{x-y}\, \mathrm{d}\lambda(y) \\
&= e^{a-x} + e^{x-b} + e^{y-x}\Big|_a^x - e^{x-y}\Big|_x^b \\
&= e^{a-x} + e^{x-b} + e^{x-x} - e^{a-x} - e^{x-b} + e^{x-x} \\
&= 2.
\end{aligned}$$

□

Therefore, the magnitude of $[a,b]$ is given by

$$|[a,b]| = \mu([a,b]) = \frac{1}{2}(\delta_a([a,b]) + \lambda([a,b]) + \delta_b([a,b])) = 1 + \frac{b-a}{2}.$$

Remark 6.7. *Fubini implies that if X and Y admit weight measures μ^X and μ^Y, then $\mu^X \times \mu^Y$ is a weight measure an $X \times Y$ with metric*

$$d^{X\times Y}((x,y),(x',y')) = d^X(x,x') + d^Y(y,y').$$

Therefore, the magnitude of $X \times Y$ is given by

$$|X \times Y| = |X| \cdot |Y|.$$

Example 6.4. *By Remark 6.7 and Example 6.3, the magnitude of a rectangle $[a,b] \times [c,d]$ with Manhattan metric*

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

is given by

$$|[a,b] \times [c,d]| = |[a,b]| \cdot |[c,d]| = \left(1 + \frac{b-a}{2}\right) \cdot \left(1 + \frac{d-c}{2}\right).$$

Proposition 6.5 (Theorem 3.4 in [16])**.** *Let (X,d) be a compact PDMS. Then*

$$|X| = \sup\left\{\frac{\mu(X)^2}{\|\mu\|_{\mathrm{W}}^2} : \mu \in \mathrm{FM}(X),\ \mu \neq 0\right\}.$$

Proof. Denote the supremum in Proposition 6.5 by

$$\kappa \coloneqq \sup\left\{\frac{\mu(X)^2}{\|\mu\|_{\mathrm{W}}^2} : \mu \in \mathrm{FM}(X),\ \mu \neq 0\right\} \in [0,\infty].$$

Let us first prove that if X admits a weight measure, then $\kappa < \infty$. Then we prove that if $\kappa < \infty$, then X admits a weight measure and $|X| = \kappa$. Together, this proves the statement.

- Suppose first that X admits a weight measure $\mu \in \mathrm{W}(X)$. Then, for all $\nu \in \mathrm{FM}(X)$ we find

$$\nu(X) = \int 1 \,\mathrm{d}\nu(y) = \int \int e^{-d(x,y)} \,\mathrm{d}\mu(x) \,\mathrm{d}\nu(y) = \langle \nu, \mu \rangle.$$

By the Cauchy-Schwarz inequality we find that

$$\nu(X)^2 = \langle \nu, \mu \rangle^2 \leq \|\nu\|_{\mathrm{W}}^2 \|\mu\|_{\mathrm{W}}^2.$$

Therefore

$$\frac{\nu(X)^2}{\|\nu\|_{\mathrm{W}}^2} \leq \|\mu\|_{\mathrm{W}}^2 = |X|.$$

- Now suppose that

$$\kappa \coloneqq \sup \left\{ \frac{\mu(X)^2}{\|\mu\|_{\mathrm{W}}^2} : \mu \in \mathrm{FM}(X),\, \mu \neq 0 \right\} < \infty.$$

The linear functional $\mu \mapsto \mu(X)$ on $(\mathrm{FM}(X), \|\cdot\|_{\mathrm{W}})$ is bounded with norm

$$\|\cdot \mapsto \cdot(X)\| = \sup \left\{ \frac{\|\mu(X)\|}{\|\mu\|_{\mathrm{W}}} : \mu \in \mathrm{FM}(X),\, \mu \neq 0 \right\} = \sqrt{\kappa}.$$

We can extend this bounded linear functional on $(\mathrm{W}, \|\cdot\|_{\mathrm{W}})$ with norm $\sqrt{\kappa}$. By Riesz representation theorem, there exists $\nu \in \mathrm{W}(X)$ such that $\|\nu\|_{\mathrm{W}} = \sqrt{\kappa}$ and

$$\forall \mu \in \mathrm{FM}(X) \colon \mu(X) = \langle \mu, \nu \rangle_{\mathrm{W}}.$$

Therefore, for all $x \in X$

$$\int_X e^{-d(x,yb)} \,\mathrm{d}\nu(y) = \int_X \int_X e^{-d(x',y)} \,\mathrm{d}\delta_x(x') \,\mathrm{d}\nu(y) = \langle \delta_x, \nu \rangle = \delta_x(X) = 1.$$

Hence $\nu \in \mathrm{W}(X)$ is a weighting of X and $|X| = \kappa$.

□

Corollary 6.1. *Suppose that (X, d) is a compact PDMS and $X' \subseteq X$, then $|X'| \leq |X|$.*

Proof. Every measure $\mu \in \mathrm{FM}(X')$ induces a measure $\nu \in \mathrm{FM}(X)$ by restriction. Hence, the statement follows from Proposition 6.5. □

6.2 Continuity of Magnitude

We aim to estimate the magnitude of infinite metric spaces by computing the magnitude of finite metric spaces that closely resemble the infinite ones. For this to work, it is crucial that the magnitudes of two similar metric spaces – meaning two metric spaces with small Gromov-Hausdorff distance – are similar. We follow [12] for this.

Theorem 6.1 (Proposition 3.1. in [12])**.** *The function*

$$\{X\colon \text{non-empty, compact PDMS}\} \to [0,\infty],\ X \mapsto |X|,$$

which assigns to each non-empty, compact PDMS its magnitude, is lower semi-continuous with respect to the Gromov-Hausdorff metric.

Let us prove the following two lemmas in preparation.

Lemma 6.2. *Let A' and B' be two finite, non-empty, compact PDMSs that satisfy $d_{\mathrm{GH}}(A', B') < \delta$. Then*

$$|B'| \geq |A'| - 6\|w\|_1^2\delta,$$

where $w \in \mathbb{R}^{A'}$ denotes the unique weighting of A'.

Proof. Since $d_{\mathrm{GH}}(A', B') < \delta$, there exists a metric space (X, d^X) and isometric embeddings $\varphi\colon A' \hookrightarrow X$, $\psi\colon B' \hookrightarrow X$ such that

$$d_{\mathrm{H}}^X(\varphi(A'), \psi(B')) \leq d_{\mathrm{GH}}(A', B') + \delta < 2\delta.$$

Therefore,

$$\sup_{a\in A'} \inf_{b\in B'} d^X(\varphi(a), \psi(b)) = \sup_{a\in A'} d^X(\varphi(a), \psi(B')) \leq d_{\mathrm{H}}^X(\varphi(A'), \psi(B')) < 2\delta.$$

Hence, for all $a \in A'$, there exists $f(a) \in B'$ such that

$$d^X(\varphi(a), \psi(f(a))) < 3\delta.$$

It follows that for all $a, a' \in A'$

$$\begin{aligned}\|d^{A'}(a,a') - d^{B'}(f(a), f(a'))\| &= \|d^X(\varphi(a), \varphi(a')) - d^X(\psi(f(a)), \psi(f(a')))\| \\ &\leq d^X(\varphi(a), \psi(f(a))) + d^X(\varphi(a'), \psi(f(a'))) \\ &< 3\delta + 3\delta = 6\delta.\end{aligned}$$

Let now $w \in \mathbb{R}^{A'}$ be the unique weighting of A'. Define $Z_{f(A')} \in \mathbb{R}^{A'\times A'}$ by

$$Z_{f(A')}(a,a') := e^{-d^{B'}(f(a),f(a'))}$$

and $u \in \mathbb{R}^{B'}$ by

$$u(b) := \sum_{a\in f^{-1}(\{b\})} w(a).$$

Then

$$\sum_{b\in B'} u(b) = \sum_{a\in A'} w(a).$$

Furthermore, if $Z_{B'} \in \mathbb{R}^{B'\times B'}$ denotes the ASM of B', then

$$\begin{aligned}u^T Z_{B'} u &= \sum_{b,b'\in B'} u(b) Z_{B'}(b,b') u(b') \\ &= \sum_{b,b'\in B'} \left(\sum_{a\in f^{-1}(\{b\})} w(a)\right) Z_{B'}(b,b') \left(\sum_{a'\in f^{-1}(\{b'\})} w(a')\right) \\ &= \sum_{a,a'\in A'} w(a) Z_{f(A')}(a,a') w(a') \\ &= w^T Z_{f(A')} w.\end{aligned}$$

Therefore we find that

$$\begin{aligned}
\|u^{\mathrm{T}} Z_{B'} u - w^{\mathrm{T}} Z_{A'} w\| &= \|w^{\mathrm{T}} \left(Z_{f(A')} - Z_{A'}\right) w\| \\
&\leq \sum_{a,a' \in A'} \|w(a)(Z_{f(A')}(a,a') - Z_{A'})(a,a')w(a')\| \\
&\leq \|Z_{f(A')} - Z_{A'}\|_\infty \sum_{a,a' \in A'} \|w(a)w(a')\| \\
&= \|Z_{f(A')} - Z_{A'}\|_\infty \left(\sum_{a \in A'} \|w(a)\|\right)^2 \\
&= \|Z_{f(A')} - Z_{A'}\|_\infty \|w\|_1^2.
\end{aligned}$$

We have

$$\begin{aligned}
\|Z_{f(A')} - Z_{A'}\|_\infty &= \sup_{a,a' \in A'} \|e^{-d^{B'}(f(a),f(a'))} - e^{-d^{A'}(a,a')}\| \\
&\leq \sup_{a,a' \in A'} \|d^{B'}(f(a), f(a')) - d^{A'}(a,a')\| < 6\delta.
\end{aligned}$$

Therefore, by Proposition 6.3

$$|B'| \geq \frac{\left(\sum_{b \in B'} u(b)\right)^2}{u^{\mathrm{T}} Z_{B'} u} \geq \frac{\left(\sum_{a \in A'} w(a)\right)^2}{w^{\mathrm{T}} Z_{A'} w + 6\delta \|w\|_1^2} \geq |A'| - 6\delta \|w\|_1^2.$$

□

Lemma 6.3. *Let A' and B be two non-empty, compact PDMSs where A' is finite such that they satisfy $d_{\mathrm{GH}}(A', B) < \delta$. Then there exists a finite subspace $B' \subseteq B$ such that $d_{\mathrm{GH}}(A', B') < 2\delta$.*

Proof. Since $d_{\mathrm{GH}}(A', B) < \delta$, there exists a metric space (X, d^X) and isometric embeddings $\varphi \colon A' \hookrightarrow X$, $\psi \colon B \hookrightarrow X$ such that

$$d_{\mathrm{H}}^X(\varphi(A'), \psi(B)) \leq d_{\mathrm{GH}}(A', B) + \delta < 2\delta.$$

In particular, we have

$$\max_{a \in A'} d^X(\varphi(a), \psi(B)) < 2\delta.$$

For all $a \in A'$ let $b_a \in B$ be such that $d^X(\varphi(a), \psi(b_a)) < 2\delta$. Set

$$B' := \{b_a \in B \colon a \in A'\}.$$

Then B' is finite and satisfies $d_{\mathrm{GH}}(A', B') < 2\delta$. □

Proof. Let X be a non-empty, compact PDMS and let us show that the magnitude function is lower semi-continuous at X. We consider the cases $|X| < \infty$ and $|X| = \infty$ separately.

- Suppose first that $|X| < \infty$:
  Let $\varepsilon > 0$. Then there exists a finite subspace $X' \subseteq X$ such that

$$|X'| \geq |X| - \frac{\varepsilon}{2}.$$

Denote by $w \in \mathbb{R}^{X'}$ the unique weighting of X' and set

$$\delta := \frac{\varepsilon}{24\|w\|_1^2} > 0.$$

Let Y be a non-empty, compact PDMS with

$$d_{\mathrm{GH}}(X, Y) < \delta.$$

By Lemma 6.3 there exists a finite $Y' \subseteq Y$ such that $d_{\mathrm{GH}}(X', Y') < 2\delta$. By Lemma 6.2 we have

$$|Y'| \geq |X'| - 12\delta\|w\|_1^2 = |X'| - \frac{\varepsilon}{2}.$$

Then we find that

$$|Y| \geq |Y'| \geq |X'| - \frac{\varepsilon}{2} \geq |X| - \varepsilon,$$

proving that magnitude function is lower semi-continuous at X.

- Now suppose that $|X| = \infty$:
  Let $K > 0$. There exists a finite subspace $X' \subseteq X$ such that $d_{\mathrm{GH}}(X, X') < \delta$ and
  $$|X'| > K + 1.$$
  Denote by $w \in \mathbb{R}^{X'}$ the unique weighting of X' and set
  $$\delta := \frac{1}{24\|w\|_1^2} > 0.$$
  Let Y be a non-empty, compact PDMS with
  $$d_{\mathrm{GH}}(X, Y) < \delta.$$
  Then $d_{\mathrm{GH}}(X', Y) < 2\delta$ and hence, by Lemma 6.3 there exists a finite $Y' \subseteq Y$ such that $d_{\mathrm{GH}}(X', Y') < 4\delta$. By Lemma 6.2, we have
  $$|Y| \geq |Y'| \geq |X'| - 24\|w\|_1^2 = K,$$
  proving that magnitude function is lower semi-continuous at X.

□

Theorem 6.2. *Let (X, d) be a non-empty, compact PDMS and let $\{X_k\}_{k\in\mathbb{N}}$ be a sequence of non-empty, compact metric subspaces $X_k \subseteq X$ such that $X_k \xrightarrow{k\to\infty} X$ in the Gromov-Hausdorff topology, that is,*

$$\lim_{k\to\infty} d_{\mathrm{GH}}(X_k, X) = 0.$$

Then the magnitudes $|X_k|$ converge to the magnitude $|X|$, that is,

$$|X| = \lim_{k\to\infty} |X_k|.$$

Proof. For all $k \in \mathbb{N}$, we have $X_k \subseteq X$ and hence $|X_k| \leq |X|$. Therefore

$$\limsup_{k\to\infty} |X_k| \leq |X|.$$

By Theorem 6.1, the magnitude function is lower semi-continuous and hence

$$\liminf_{k\to\infty} |X_k| \geq |X|.$$

Hence we can conclude

$$\lim_{k\to\infty} |X_k| = |X|.$$

□

This means that we can compute the magnitude of a finite sample to estimate the magnitude of the underlying space.

6.3 Magnitude Dimension

After introducing magnitude for metric spaces and proving some properties, we now focus on the magnitude function which describes how the magnitude of a metric space changes as we scale the space. This leads to the definition of the magnitude dimension as the growth rate of the magnitude function. We also relate the magnitude dimension to the Minkowski dimension and prove that they agree for compact subsets of $\mathbb{R}^n$. For this, we follow [16] closely.
For the remainder of this section, let (X, d) be a compact metric space.

Definition 6.12. *The* **maximum diversity** $|X|_+$ *of* X *is defined as*

$$|X|_+ \coloneqq \sup \left\{ \frac{\mu(X)^2}{\|\mu\|_{\mathrm{W}}^2} : \mu \in \mathrm{M}_+(X),\, \mu \neq 0 \right\}.$$

Proposition 6.6. *We have*

$$|X|_+ = \sup_{\mu \in \mathcal{P}(X)} \left(\int_X \int_X e^{-d(x,y)} \,\mathrm{d}\mu(y) \,\mathrm{d}\mu(x) \right)^{-1},$$

where $\mathcal{P}(X)$ *denotes the space of Borel probability measures on* X.

Proof. For any $\lambda > 0$ and $\mu \in \mathrm{M}_+(X)$ we have

$$\frac{(\lambda\mu(X))^2}{\|\lambda\mu\|_{\mathrm{W}}^2} = \frac{\mu(X)}{\|\mu\|_{\mathrm{W}}^2}.$$

Therefore, we can simply restrict the supremum in Definition 6.12 to all measures $\mu \in \mathrm{M}_+(X)$ with total measure $\mu(X) = 1$, which is exactly $\mathcal{P}(X)$. □

Definition 6.13. *For* $t > 0$ *denote by* tX *the metric space* X *scaled by* t*, that is,* tX *has the same points as* X *and the distance in* tX *is given by*

$$\forall x, y \in X \colon d^{tX}(x, y) = t \cdot d^X(x, y).$$

We call X **stably positive definite**, *if* tX *is positive definite for all* $t > 0$.

If X is stably positive definite, then the magnitude $|tX|$ exists for every scale $t > 0$, and we can define the magnitude function.

Definition 6.14. *Suppose (X, d) is stably positive definite. The* **magnitude function** *on X is*

$$(0, \infty) \to [0, \infty],\ t \mapsto |tX|.$$

Definition 6.15. *The* **instantaneous magnitude dimension** *of X at some scale $s > 0$ is defined as the growth rate of the magnitude function at s*

$$\dim_{\text{mag}}^{\text{inst}}(X, s) := \frac{\mathrm{d}\log(|tX|)}{\mathrm{d}\log(t)}\Big|_{t=s} = \frac{t}{|tX|}\frac{\mathrm{d}|tX|}{\mathrm{d}t}\Big|_{t=s}.$$

This is precisely the slope of the magnitude function over the scale parameter when viewed in a log-log plot. The instantaneous magnitude dimension $\dim_{\text{mag}}^{\text{inst}}(X, s)$ tells us the 'dimension' of X when viewed at scale s.

Example 6.5. *Consider the very long and thin rectangle*

$$R := [0, 2] \times [0, 200000]$$

with Manhattan metric

$$d((x, y), (x', y')) = \|x' - x\| + \|y' - y\|.$$

The rectangle can appear as 0-, 1- or 2-dimensional depending on the scale it which it is viewed. This is nicely depicted in Figure 6.1 By Example 6.4, the

(a) Small scale. (b) Medium scale. (c) Large scale.

Figure 6.1: A long, thin rectangle at different scales.

magnitude function of R is given by

$$(0, \infty) \to [0, \infty],\ t \mapsto |tR| = \left(1 + \frac{2t}{2}\right) \cdot \left(1 + \frac{200000t}{2}\right).$$

Figure 6.2 shows the magnitude function in a log-log plot. The instantaneous magnitude dimension of R is plotted in Figure 6.3. We can see that the long, thin rectangle has instantaneous magnitude dimension 0 at small scale, 1 at medium scale and 2 at large scale.

If we do not know the magnitude of an infinite metric space, we may approximate its magnitude via a finite subspace. We can then compute the instantaneous magnitude dimension of the finite approximation as an estimate for the dimension of the infinite space.

Example 6.6. *Denote by X the space of 10000 points from S_5 depicted in Figure 6.4 as an approximation of the Sierpinski triangle. Denote by $Z_{tX} \in \mathbb{R}^{X \times X}$ the ASM of tX. For each scale t we find a weighting w^t of tX by solving the linear system $Z_{tX} \cdot w^t = \mathbb{1}$ and compute the magnitude $|tX|$. Figure 6.5 shows the magnitude function of X in a log-log plot. The instantaneous magnitude dimension of X is plotted in Figure 6.6. We see that at low scale, X looks like*

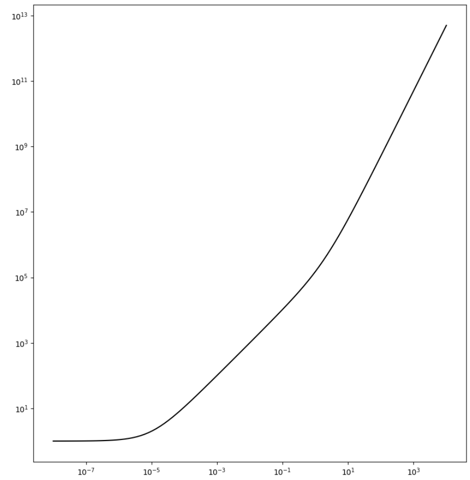


Figure 6.2: Magnitude function of a long, thin rectangle.

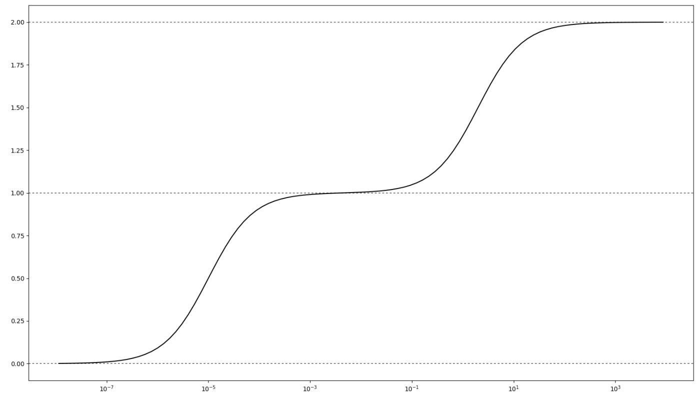


Figure 6.3: Instantaneous magnitude dimension of a long, thin rectangle.

a single point and hence has an instantaneous magnitude dimension close to 0. *At medium scale, X looks very close to the Sierpinski triangle and reaches its maximum instantaneous magnitude dimension of approximately* 1.5396. *This is very close to the Minkowski dimension of the Sierpinski triangle*

$$\dim_{\mathrm{mink}}(\mathcal{S}) = \frac{\log(3)}{\log(2)} \approx 1.5850.$$

When viewed at larges scale (see Figure 6.7) the individual points of X become visible and its instantaneous magnitude dimension decreases to 0.

The magnitude dimension is defied as the asymptotic growth rate of the magnitude function.

Definition 6.16. *Let (X, d) be stably positive definite. The* **lower magnitude dimension** *of X is*

$$\underline{\dim}_{\mathrm{mag}}(X) \coloneqq \liminf_{t\to\infty} \frac{\log(|tX|)}{\log(t)}$$

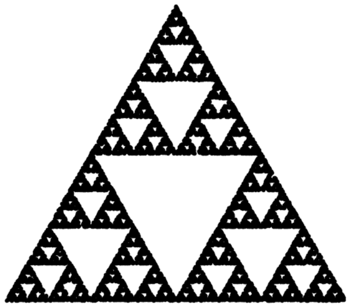


Figure 6.4: 10000 points from S_5.

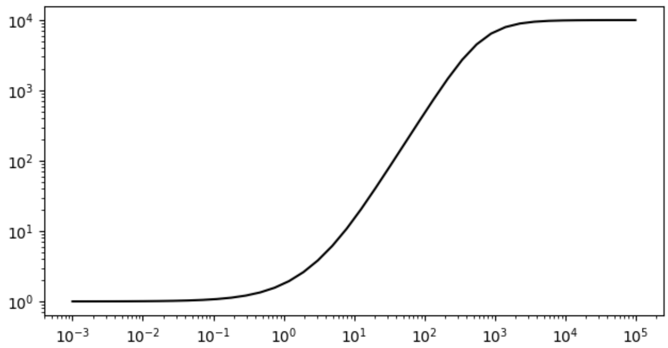


Figure 6.5: Magnitude function.

and the **upper magnitude dimension** *of X is*

$$\overline{\dim}_{\mathrm{mag}}(X) := \limsup_{t\to\infty} \frac{\log(|tX|)}{\log(t)}.$$

If $\underline{\dim}_{\mathrm{mag}}(X) = \overline{\dim}_{\mathrm{mag}}(X)$, the **magnitude dimension** *of X exists and is*

$$\dim_{\mathrm{mag}}(X) := \lim_{t\to\infty} \frac{\log(|tX|)}{\log(t)}.$$

Definition 6.17. *The* **lower** *and* **upper diversity dimension** *of X are defined analogously with the maximum diversity $|tX|_+$ instead of the magnitude $|tX|$:*

$$\underline{\dim}_{\mathrm{div}}(X) := \liminf_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)}, \quad \overline{\dim}_{\mathrm{div}}(X) := \limsup_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)}.$$

If $\underline{\dim}_{\mathrm{div}}(X) = \overline{\dim}_{\mathrm{div}}(X)$, the **diversity dimension** *exists and is*

$$\dim_{\mathrm{div}}(X) := \lim_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)}.$$

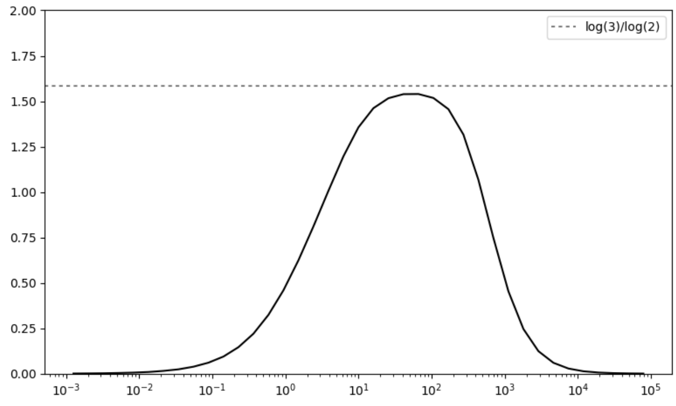


Figure 6.6: Instantaneous magnitude dimension.

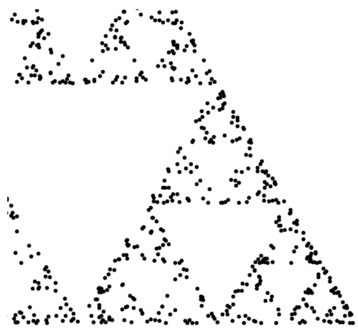


Figure 6.7: Part of X at large scale.

Remark 6.8. *We have*

$$|X|_+ \leq |X|$$

and so if (X, d) *is stably positive definite, then*

$$\underline{\dim}_{\text{div}}(X) \leq \underline{\dim}_{\text{mag}}(X), \quad \overline{\dim}_{\text{div}}(X) \leq \overline{\dim}_{\text{mag}}(X).$$

Remark 6.9 (Corollary 6.2 in [16])**.** *For all* $n \in \mathbb{N}$ *there exists* $\kappa_n \in \mathbb{R}$ *such that*

$$\forall X \subseteq \mathbb{R}^n \textit{ compact}\colon |X| \leq \kappa_n |X|_+.$$

We refer to [16] for a proof involving results from potential theory from [1].

Theorem 6.3 (Theorem 7.1 in [16])**.** *For all compact metric spaces* (X, d) *we have that*

$$\underline{\dim}_{\text{div}}(X) = \underline{\dim}_{\text{mink}}(X), \quad \overline{\dim}_{\text{div}}(X) = \overline{\dim}_{\text{mink}}(X).$$

Hence $\dim_{\text{div}}(X)$ *is defined if and only if* $\dim_{\text{mink}}(X)$ *is defined, and in that case*

$$\dim_{\text{div}}(X) = \dim_{\text{mink}}(X).$$

Proof. Let us first prove that

$$\underline{\dim}_{\text{div}}(X) \leq \underline{\dim}_{\text{mink}}(X), \quad \overline{\dim}_{\text{div}}(X) \leq \overline{\dim}_{\text{mink}}(X).$$

Let $\varepsilon > 0$, $t > 0$ and $\mu \in \mathcal{P}(X)$. For all $x \in X$ we find

$$\begin{aligned}
\int_X e^{-td(x,y)} \mathrm{d}\mu(y) &\geq \int_{B(x,\varepsilon)} e^{-td(x,y)} \mathrm{d}\mu(y) \\
&\geq \int_{B(x,\varepsilon)} e^{-t\varepsilon} \mathrm{d}\mu(y) \\
&= e^{-t\varepsilon} \mu(B(x,\varepsilon)).
\end{aligned}$$

Therefore, by applying Jensen's inequality, we find that

$$\begin{aligned}
\left(\int_X \int_X e^{-td(x,y)} \mathrm{d}\mu(y) \mathrm{d}\mu(x) \right)^{-1} &\leq \left(\int_X e^{-t\varepsilon} \mu(B(a,\varepsilon)) \mathrm{d}\mu(x) \right)^{-1} \\
&\leq \int_X \left(e^{-t\varepsilon} \mu(B(a,\varepsilon)) \right)^{-1} \mathrm{d}\mu(x) \\
&\leq e^{t\varepsilon} \int_X \frac{1}{\mu(B(a,\varepsilon))} \mathrm{d}\mu(x).
\end{aligned}$$

Denote by $N \coloneqq N(X, \frac{\varepsilon}{2})$ the $\left(\frac{\varepsilon}{2}\right)$-covering number of X and let $\{x_1, \ldots, x_N\} \subseteq X$ such that we can write

$$X = \bigcup_{j=1}^{N} B\left(x_j, \frac{\varepsilon}{2}\right).$$

If $x \in B\left(x_j, \frac{\varepsilon}{2}\right)$ for some j, then $B\left(x_j, \frac{\varepsilon}{2}\right) \subseteq B(x, \varepsilon)$ and hence

$$\mu\left(B\left(x_j, \frac{\varepsilon}{2}\right)\right) \leq \mu(B(x,\varepsilon)).$$

Therefore

$$\begin{aligned}
\int_X \frac{1}{\mu(B(x,\varepsilon))} \mathrm{d}\mu(x) &\leq \sum_{j=1}^{N} \int_{B(x_j, \frac{\varepsilon}{2})} \frac{1}{\mu(B(x,\varepsilon))} \mathrm{d}\mu(x) \\
&\leq \sum_{j=1}^{N} \int_{B(x_j, \frac{\varepsilon}{2})} \frac{1}{\mu(B(x_j, \frac{\varepsilon}{2}))} \mathrm{d}\mu(x) \\
&= \sum_{j=1}^{N} \frac{\mu(B(x_j, \frac{\varepsilon}{2}))}{\mu(B(x_j, \frac{\varepsilon}{2}))} = N.
\end{aligned}$$

Putting everything together, we find that

$$\begin{aligned}
|tX|_+ &= \sup_{\mu \in \mathcal{P}(X)} \left(\int_X \int_X e^{-td(x,y)} \mathrm{d}\mu(y) \mathrm{d}\mu(x) \right)^{-1} \\
&\leq \sup_{\mu \in \mathcal{P}(X)} e^{t\varepsilon} \int_X \frac{1}{\mu(B(x,\varepsilon))} \mathrm{d}\mu(x) \\
&\leq \sup_{\mu \in \mathcal{P}(X)} N e^{t\varepsilon} = N\left(X, \frac{\varepsilon}{2}\right) e^{t\varepsilon}.
\end{aligned}$$

If we set $\varepsilon = \frac{2}{t}$, we get

$$\log(|tX|_+) \leq \log\left(N\left(X, \frac{\varepsilon}{2}\right)e^{t\varepsilon}\right) \leq \log\left(N\left(X, \frac{\varepsilon}{2}\right)\right).$$

Therefore we can conclude

$$\underline{\dim}_{\mathrm{div}}(X) = \liminf_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)} \leq \liminf_{\varepsilon\to 0^+} \frac{\log(N(X, \frac{\varepsilon}{2})}{\log(\frac{2}{\varepsilon}))} = \underline{\dim}_{\mathrm{mink}}(X)$$

and analogously

$$\overline{\dim}_{\mathrm{div}}(X) = \limsup_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)} \leq \limsup_{\varepsilon\to 0^+} \frac{\log(N(X, \frac{\varepsilon}{2})}{\log(\frac{2}{\varepsilon}))} = \overline{\dim}_{\mathrm{mink}}(X).$$

Let us now prove that

$$\underline{\dim}_{\mathrm{mink}}(X) \leq \underline{\dim}_{\mathrm{div}}(X), \quad \overline{\dim}_{\mathrm{mink}}(X) \leq \overline{\dim}_{\mathrm{div}}(X).$$

Let $\varepsilon > 0$ and $t > 0$. Denote by $M := M(X, \varepsilon)$ the ε-packing number of X and let $\{x_1, \dots, x_M\} \subseteq X$ such that $\{B(x_j, \varepsilon)\}_{j=1}^M$ are disjoint. Consider the measure

$$\mu := \frac{1}{M}\sum_{j=1}^M \delta_{x_j} \in \mathcal{P}(X).$$

Since $\{B(x_j, \varepsilon)\}_{j=1}^M$ are disjoint, every $x \in X$ can be in at most one such ball. Therefore

$$\int_X e^{-td(x,y)}\mathrm{d}\mu(y) = \frac{1}{M}\sum_{j=1}^M e^{-td(x,x_j)} \leq \frac{1}{M}\left(1 + (M-1)e^{-t\varepsilon}\right) \leq \frac{1}{M} + e^{-t\varepsilon}.$$

Hence we find that

$$\begin{aligned}
\frac{1}{|tX|_+} &= \left(\sup_{\nu\in\mathcal{P}(X)}\left(\int_X\int_X e^{-td(x,y)}\mathrm{d}\nu(y)\mathrm{d}\nu(x)\right)^{-1}\right)^{-1} \\
&= \inf_{\nu\in\mathcal{P}(X)} \int_X\int_X e^{-td(x,y)}\mathrm{d}\nu(y)\mathrm{d}\nu(x) \\
&\leq \int_X\int_X e^{-td(x,y)}\mathrm{d}\mu(y)\mathrm{d}\mu(x) \\
&\leq \int_X \frac{1}{M} + e^{-t\varepsilon}\mathrm{d}\mu(x) \\
&= \frac{1}{M}\sum_{j=1}^M\left(\frac{1}{M} + e^{-t\varepsilon}\right) \\
&= \frac{1}{M(X,\varepsilon)} + e^{-t\varepsilon}.
\end{aligned}$$

If we set $\varepsilon = \frac{\log(2|tX|_+)}{t}$, then $e^{-t\varepsilon} = \frac{1}{2|tX|_+}$ and hence

$$M(X, \varepsilon) \leq 2|tX|_+.$$

Therefore

$$\frac{\log(M(X,\varepsilon))}{\log(\frac{1}{\varepsilon})} \leq \frac{\log(2|tX|_+)}{\log(\frac{1}{\varepsilon})} = \frac{\log(|tX|_+)}{\log(t)} \frac{\log(t)}{\log(\frac{1}{\varepsilon})} + \frac{\log(2)}{\log(\frac{1}{\varepsilon})}.$$

If $\underline{\dim}_{\text{div}}(X) = \infty$, then $\overline{\dim}_{\text{div}}(X) = \infty$ and so

$$\underline{\dim}_{\text{mink}}(X) \leq \underline{\dim}_{\text{div}}(X), \quad \overline{\dim}_{\text{mink}}(X) \leq \overline{\dim}_{\text{div}}(X)$$

hold trivially. So we may assume that $\underline{\dim}_{\text{div}}(X) < \infty$. This implies that $\varepsilon(t) \xrightarrow{t\to\infty} 0$. Indeed if $\varepsilon(t) \geq c$ for some constant $c > 0$, then $\log(2|tX|_+) \geq ct$ which leads to the contradiction

$$\begin{aligned}\underline{\dim}_{\text{div}}(X) &= \liminf_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)} \\ &= \liminf_{t\to\infty} \frac{\log(2|tX|_+) - \log(2)}{\log(t)} \\ &\geq \liminf_{t\to\infty} \frac{ct - \log(2)}{\log(t)} = \infty.\end{aligned}$$

Therefore, we find that

$$\frac{\log(2)}{\log(\frac{1}{\varepsilon})} = o(1) \quad \text{as } t \to \infty.$$

Furthermore we have

$$\frac{\log\left(\frac{1}{\varepsilon}\right)}{\log(t)} = \frac{\log\left(\frac{t}{\log(2|tX|_+)}\right)}{\log(t)} = \frac{\log(t) - \log(\log(2|tX|_+))}{\log(t)} = 1 - \frac{\log(\log(2|tX|_+))}{\log(t)}.$$

We find that

$$\begin{aligned}\frac{\log(\log(2|tX|_+)}{\log(t)} &= \frac{\log(\log(2|tX|_+) - \log(\log(t))}{\log(t)} + \frac{\log(\log(t))}{\log(t)} \\ &= \frac{\log\left(\frac{\log(2|tX|_+)}{\log(t)}\right)}{\log(t)} + o(1) \quad \text{as } t \to \infty.\end{aligned}$$

Therefore, under the assumption $K \coloneqq \underline{\dim}_{\text{div}}(X) < \infty$, we find that

$$\liminf_{t\to\infty} \frac{\log(\log(2|tX|_+)}{\log(t)} = \liminf_{t\to\infty} \frac{\log(K)}{\log(t)} = 0.$$

Hence we can conclude

$$\begin{aligned}\underline{\dim}_{\text{mink}}(X) &= \liminf_{\varepsilon\to 0^+} \frac{\log(M(X,\varepsilon))}{\log(\frac{1}{\varepsilon})} \\ &\leq \liminf_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)} \frac{\log(t)}{\log(\frac{1}{\varepsilon})} + \frac{\log(2)}{\log(\frac{1}{\varepsilon})} \\ &\leq \liminf_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)} \\ &= \underline{\dim}_{\text{div}}(X).\end{aligned}$$

Analogously, under the assumption $K' \coloneqq \overline{\dim}_{\mathrm{div}}(X) < \infty$, we find that

$$\limsup_{t\to\infty} \frac{\log(\log(2|tX|_+))}{\log(t)} = \limsup_{t\to\infty} \frac{\log(K')}{\log(t)} = 0.$$

Hence we can conclude

$$\begin{aligned}
\overline{\dim}_{\mathrm{mink}}(X) &= \limsup_{\varepsilon\to 0^+} \frac{\log(M(X,\varepsilon))}{\log(\frac{1}{\varepsilon})} \\
&\leq \limsup_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)} \frac{\log(t)}{\log(\frac{1}{\varepsilon})} + \frac{\log(2)}{\log(\frac{1}{\varepsilon})} \\
&\leq \limsup_{t\to\infty} \frac{\log(|tX|_+)}{\log(t)} \\
&= \overline{\dim}_{\mathrm{div}}(X).
\end{aligned}$$

□

Corollary 6.2. *Let (X,d) be a compact stably positive definite metric space. Then $\underline{\dim}_{\mathrm{mink}}(X) \leq \underline{\dim}_{\mathrm{mag}}(X)$ and $\overline{\dim}_{\mathrm{mink}}(X) \leq \overline{\dim}_{\mathrm{mag}}(X)$.*

Proof. This follows immediately from Theorem 6.3 and Remark 6.8. □

Theorem 6.4. *Let $X \subseteq \mathbb{R}^n$ be compact. Then*

$$\underline{\dim}_{\mathrm{mag}}(X) = \underline{\dim}_{\mathrm{mink}}(X), \quad \overline{\dim}_{\mathrm{mag}}(X) = \overline{\dim}_{\mathrm{mink}}(X).$$

Hence $\dim_{\mathrm{mag}}(X)$ is defined if and only if $\dim_{\mathrm{mink}}(X)$ is defined, and in that case

$$\dim_{\mathrm{mag}}(X) = \dim_{\mathrm{mink}}(X).$$

Proof. Let $X \subseteq \mathbb{R}^n$ be compact. By Remark 6.8 and Remark 6.9 we find

$$|tX|_+ \leq |tX| \leq \kappa_n |tX|_+.$$

Therefore

$$\frac{\log(|tX|_+)}{\log(t)} \leq \frac{\log(|tX|)}{\log(t)} \leq \frac{\log(\kappa_n) + \log(|tX|_+)}{\log(t)} = \frac{\log(\kappa_n)}{\log(t)} + \frac{\log(|tX|_+)}{\log(t)}.$$

Letting $t \to \infty$ we find

$$\underline{\dim}_{\mathrm{div}}(X) \leq \underline{\dim}_{\mathrm{mag}}(X) \leq \underline{\dim}_{\mathrm{div}}(X),$$

and

$$\overline{\dim}_{\mathrm{div}}(X) \leq \overline{\dim}_{\mathrm{mag}}(X) \leq \overline{\dim}_{\mathrm{div}}(X),$$

proving equality. □

6.4 Spread Dimension

In this section we briefly touch on spread, another measurement of size similar to the magnitude, but in some ways better behaved and easier to compute. We define spread for finite metric spaces and analyse how it changes as we scale the space. This leads to the definition of the instantaneous spread dimension. We refer to [21] for more details on the matter.
Let (X, d) denote a finite metric space.

Definition 6.18. *The* **spread** *of* (X, d) *is defined as*

$$E_0(X) := \sum_{x \in X} \frac{1}{\sum_{y \in X} Z_X(x, y)},$$

where Z_X *denotes the ASM of* X.

Example 6.7. *Consider the space* X *from Example* 6.1*,* 2 *points at distance* $d > 0$*. The spread of* X *is given by*

$$E_0(X) = \frac{2}{1 + e^{-d}}.$$

Similar to the magnitude function, we can study how the spread of a metric space changes with scaling.

Definition 6.19. *The* **spread function** *on* X *is*

$$(0, \infty) \to [0, \infty],\ t \mapsto E_0(tX).$$

Similar to the instantaneous magnitude dimension, we can study the growth rate of the spread function.

Definition 6.20. *The* **instantaneous spread dimension** *of* X *as some scale* $s > 0$ *is*

$$\dim^{\text{inst}}_{\text{spread}}(X, s) := \frac{\mathrm{d}\log(E_0(tX))}{\mathrm{d}\log(t)}\Big|_{t=s} = \frac{t}{E_0(tX)} \frac{\mathrm{d}E_0(tX)}{\mathrm{d}t}\Big|_{t=s}.$$

Since the formula for spread is in a closed form, we can simply compute the derivative. For $x, y \in X$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t} Z_{tX}(x, y) = \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-td(x,y)}\right) = -d(x, y)e^{-td(x,y)} = -d(x, y)Z_{tX}(x, y).$$

Therefore

$$\begin{aligned}
\frac{\mathrm{d}E_0(tX)}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{x \in X} \frac{1}{\sum_{y \in X} Z_{tX}(x, y)}\right) \\
&= -\sum_{x \in X} \frac{\frac{\mathrm{d}}{\mathrm{d}t}\left(\sum_{y \in X} Z_{tX}(x, y)\right)}{\left(\sum_{y \in X} Z_{tX}(x, y)\right)^2} \\
&= \sum_{x \in X} \frac{\sum_{y \in X} d(x, y) Z_{tX}(x, y)}{\left(\sum_{y \in X} Z_{tX}(x, y)\right)^2}.
\end{aligned}$$

Hence the instantaneous spread dimension is given by

$$\dim^{\text{inst}}_{\text{spread}}(X,s) = \frac{s}{E_0(sX)} \sum_{x\in X} \frac{\sum_{y\in X} d(x,y) Z_{sX}(x,y)}{\left(\sum_{y\in X} Z_{sX}(x,y)\right)^2}$$

Example 6.8. *Let us compute the instantaneous spread dimension for an approximation of the Sierpinski triangle. Denote by* X *the same* 10000 *points on* S_5 *as in Example* 6.6. *Figures 6.8 and 6.9 show the spread function on* X *in a log-log plot and the instantaneous spread dimension of* X *respectively. At*

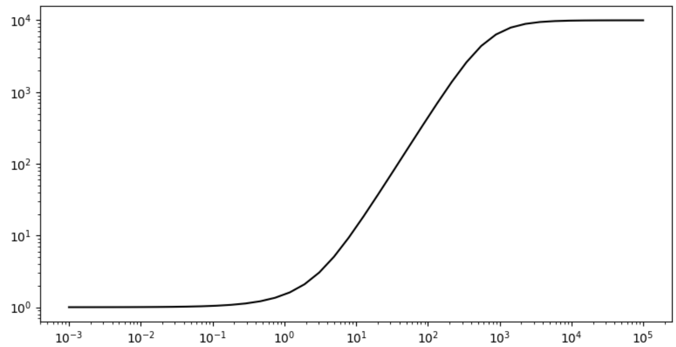


Figure 6.8: Spread function.

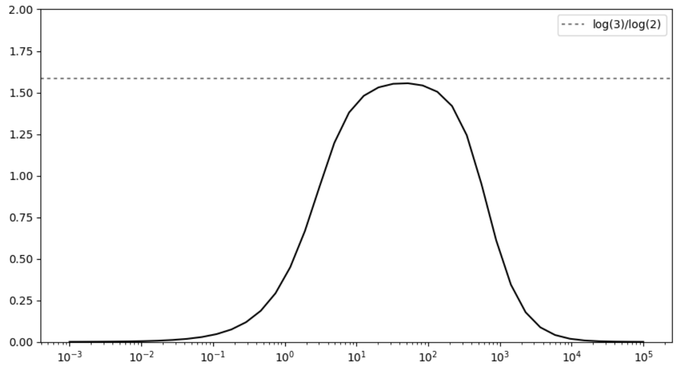


Figure 6.9: Instantaneous spread dimension.

medium scale, where X *looks very similar to the Sierpinski triangle, the instantaneous spread dimension reaches is maximum of approximately* 1.5557.

Chapter 7

Experiments

In this chapter we estimate the persistent homology and magnitude dimension of several fractals and compare them to the Minkowski and Hausdorff dimension. The python code used for all the computations is available on GitHub [10].

7.1 Cantor Dust

Construction

The Cantor dust $\mathcal{D}$ can be constructed as follows: Start with a square D_0. Now subdivide the square into nine equal squares that have side length $\frac{1}{3}$ of the original square. Remove all except the four squares in the corners. Continue this process iteratively for each of the remaining squares. The Cantor dust $\mathcal{D}$ consists exactly of the points that remain after continuing this process infinitely, that is,

$$\mathcal{D} \coloneqq \bigcap_{n=0}^{\infty} D_n.$$

Figure 7.1 shows the first six sets $D_0, \ldots, D_5$ in the construction.

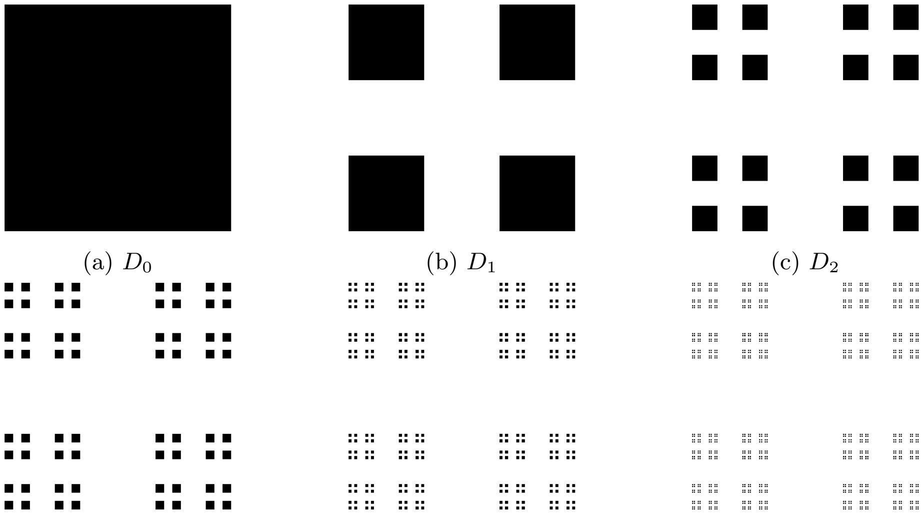
Hausdorff Dimension

Let us compute the Hausdorff dimension of the Cantor dust using Theorem 4.1. Consider $\mathcal{D}$ as subset of $[0,1]^2$ so that the maps

$$f_1, f_2, f_3, f_4 \coloneqq [0,1]^2 \to [0,1]^2$$

given by

$$\begin{aligned}
f_1(x,y) &\coloneqq \frac{1}{3}(x,y), \\
f_2(x,y) &\coloneqq \frac{1}{3}(x,y) + \frac{2}{3}(1,0), \\
f_3(x,y) &\coloneqq \frac{1}{3}(x,y) + \frac{2}{3}(0,1), \\
f_4(x,y) &\coloneqq \frac{1}{3}(x,y) + \frac{2}{3}(1,1),
\end{aligned}$$



(a) D_0 (b) D_1 (c) D_2

(d) D_3 (e) D_4 (f) D_5

Figure 7.1: Construction of the Cantor dust.

are contracting similarities with ratios $c_1, c_2, c_3, c_4 = \frac{1}{3}$. They satisfy open set condition for $(0,1)^2$ as

$$\begin{aligned}(0,1)^2 &\subseteq \left(0,\frac{1}{3}\right)^2 \sqcup \left[\left(\frac{2}{3},1\right) \times \left(0,\frac{1}{3}\right)\right] \sqcup \left[\left(0,\frac{1}{3}\right) \times \left(\frac{2}{3},1\right)\right] \sqcup \left(\frac{2}{3},1\right)^2 \\ &= S_1((0,1)^2) \sqcup S_2((0,1)^2) \sqcup S_3((0,1)^2) \sqcup S_4((0,1)^2).\end{aligned}$$

The Cantor dust is the attractor

$$\mathcal{D} = S_1(\mathcal{D}) \cup S_2(\mathcal{D}) \cup S_3(\mathcal{D}) \cup S_4(\mathcal{D}).$$

By Theorem 4.1, the Hausdorff dimension $s = \dim_{\mathrm{H}}(\mathcal{D})$ is given by

$$\left(\frac{1}{3}\right)^s + \left(\frac{1}{3}\right)^s + \left(\frac{1}{3}\right)^s + \left(\frac{1}{3}\right)^s = 1.$$

Solving for s yields

$$\dim_{\mathrm{mink}}(\mathcal{D}) = \dim_{\mathrm{H}}(\mathcal{D}) = \frac{\log(4)}{\log(3)} \approx 1.2619.$$

Persistent Homology Dimension

Let us estimate the persistent homology dimension of the uniform distribution μ on the Cantor dust. We approximate the Cantor dust by taking a uniform sample of up to 10000 points from D_5 in the construction above. Figure 7.2 shows the log-log plot of L_0 for the samples. The asymptotic slope is approximately

$$\alpha \approx 0.1903,$$

which leads to an approximate persistent homology dimension of

$$\dim_{\mathrm{PH}}(\mu) \approx \frac{1}{1-\alpha} \approx 1.2350.$$

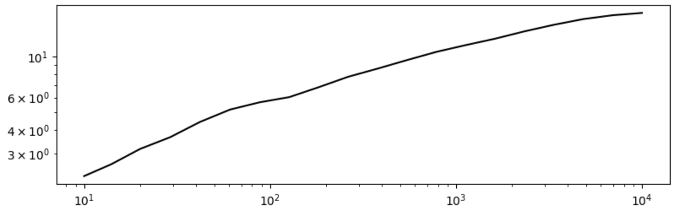


Figure 7.2: L_0 for samples of the Cantor dust.

Magnitude Dimension

To estimate the magnitude dimension of the Cantor dust, we approximate $\mathcal{D}$ by a finite set X of 10000 points as depicted in Figure 7.3. Figures 7.4a and

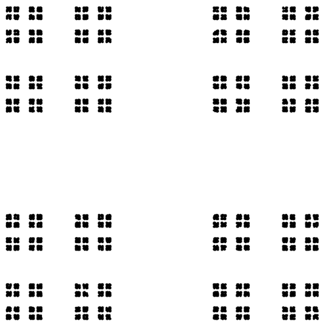


Figure 7.3: 10000 points from the Cantor dust.

7.4b shows the magnitude function and instantaneous magnitude dimension for X. At medium scale, where X looks similar to the Cantor dust, it reaches its Maximum instantaneous magnitude dimension of approximately

$$\dim_{\text{mag}}(\mathcal{D}) \approx 1.2573.$$

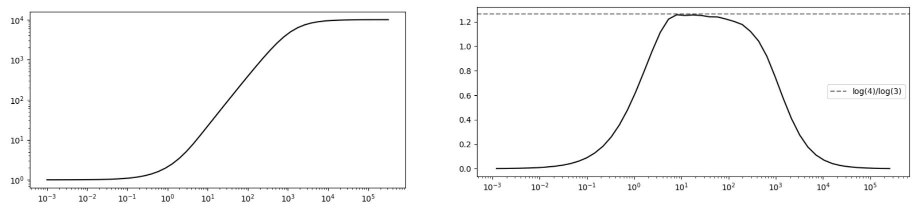
7.2 Sierpinski Triangle

We showed the construction of the Sierpinski triangle $\mathcal{S}$ in Example 3.2. In Example 5.4 we approximated the persistent homology dimension of the uniform distribution μ on $\mathcal{S}$ and found

$$\dim_{\text{PH}}(\mu) \approx 1.4679.$$

We approximated the magnitude dimension of $\mathcal{S}$ in Example 6.6

$$\dim_{\text{mag}}(\mathcal{S}) \approx 1.5396.$$



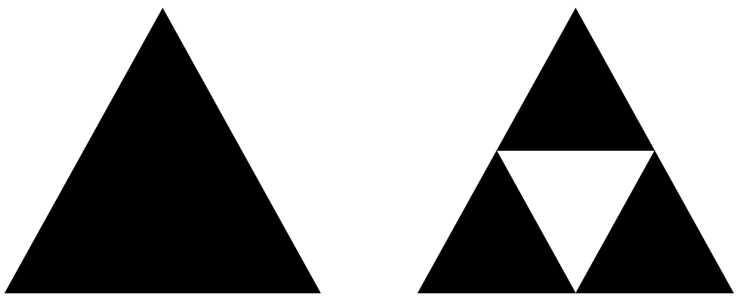
(b) Instantaneous magnitude dimension.

(a) Magnitude function.

Figure 7.4: Magnitude function and instantaneous magnitude dimension.

Hausdorff Dimension

The three maps that send the S_0 to any one of the triangles in S_1 are contracting similarities with ratio $\frac{1}{2}$ (see Figure 7.5). The open set condition is satisfied for



(a) S_0

(b) S_1

Figure 7.5: S_0 and S_1 from the construction of the Sierpinski triangle.

the interior of S_0 and so by Theorem 4.1, the Hausdorff dimension $s := \dim_{\mathrm{H}}(\mathcal{S})$ is given by

$$3 \cdot \left(\frac{1}{2}\right)^s = 1.$$

Solving for s yields

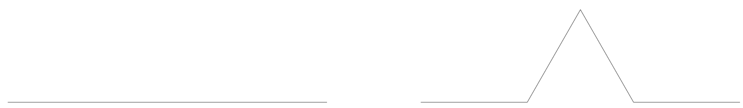
$$\dim_{\mathrm{mink}}(\mathcal{S}) = \dim_{\mathrm{H}}(\mathcal{S}) = \frac{\log(3)}{\log(2)} \approx 1.5850.$$

7.3 Koch Snowflake

We covered the construction of the Koch curve and Koch snowflake in Example 3.3. Since the Koch snowflake is made up of three copies of the Koch curve, they have equal dimensions.

Hausdorff Dimension

The four maps which send K_0^C to any one of the line segments in K_1^C are contracting similarities with ratio $\frac{1}{3}$ (see Figure 7.6). The open set condition is



(a) K_0^C (b) K_1^C

Figure 7.6: K_0^C and K_1^C from the construction of the Koch curve.

satisfied for the interior of K_0^C and so by Theorem 4.1, the Hausdorff dimension $s = \dim_{\mathrm{H}}(\mathcal{K}^C)$ is given by

$$4 \cdot \left(\frac{1}{3}\right)^s = 1.$$

Solving for s yields

$$\dim_{\mathrm{mink}}(\mathcal{K}^C) = \dim_{\mathrm{H}}(\mathcal{K}^C) = \frac{\log(4)}{\log(3)} \approx 1.2619.$$

Persistent Homology Dimension

In order to estimate the persistent homology dimension of the uniform distribution μ on the Koch curve, we samples of up to 10000 points from K_5^C as an approximation. Figure 7.7 show the log-log plot of L_0 for these samples. The

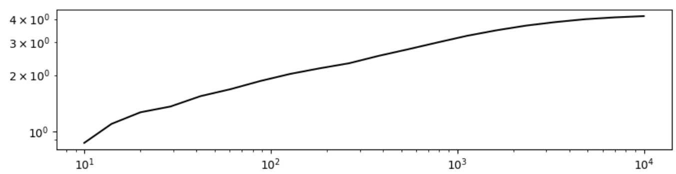


Figure 7.7: L_0 for samples of the Koch curve.

asymptotic slope is approximately

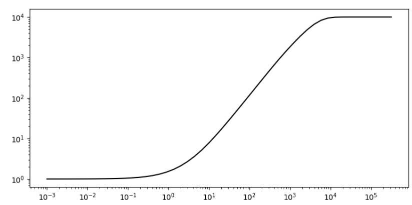
$$\alpha \approx 0.0803.$$

This yields an approximate persistent homology dimension of

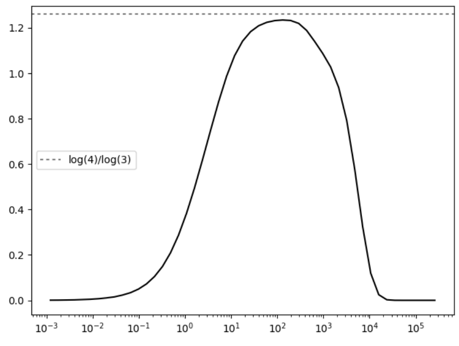
$$\dim^0_{\mathrm{PH}}(\mu) \approx 1.0873.$$

Magnitude Dimension

Consider 10000 points on K_5^C as an approximation of the Koch curve $\mathcal{K}^C$. Figure 7.8a shows the magnitude function in a log-log plot. The instantaneous magnitude dimension in depicted in Figure 7.8b. The maximum instantaneous



(a) Magnitude function.



(b) Instantaneous magnitude dimension.

Figure 7.8: Magnitude function and instantaneous magnitude dimension.

magnitude dimension is approximately

$$\dim_{\mathrm{mag}}(\mathcal{K}^C) \approx 1.2348.$$

Bibliography

[1] D.R. Adams and L.I. Hedberg. ***Function Spaces and Potential Theory***. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2012. ISBN: 9783662032824. URL: https://books.google.ch/books?id=HobqCAAAQBAJ.

[2] Henry Adams et al. "A Fractal Dimension for Measures via Persistent Homology". In: ***Abel Symposia***. Springer International Publishing, 2020, 1–31. ISBN: 9783030434083. DOI: 10.1007/978-3-030-43408-3_1. URL: http://dx.doi.org/10.1007/978-3-030-43408-3_1.

[3] Gunnar Carlsson. "Topology and Data". In: ***Bulletin of The American Mathematical Society - BULL AMER MATH SOC*** 46 (Apr. 2009), pp. 255–308. DOI: 10.1090/S0273-0979-09-01249-X.

[4] Thomas H Cormen et al. ***Introduction to algorithms***. MIT press, 2009.

[5] Pawel Dlotko. "Persistence representations". In: ***GUDHI User and Reference Manual***. GUDHI Editorial Board, 2017. URL: http://gudhi.gforge.inria.fr/doc/latest/group___persistence__representations.html.

[6] G.A. Edgar. ***Measure, Topology, and Fractal Geometry***. Undergraduate Texts in Mathematics. Springer New York, 2013. ISBN: 9781475741346. URL: https://books.google.ch/books?id=UJ3lBwAAQBAJ.

[7] K.J. Falconer. ***Fractal Geometry: Mathematical Foundations and Applications***. Fractal Geometry: Mathematical Foundations and Applications. Wiley, 2003. ISBN: 9780470848623. URL: https://books.google.ch/books?id=8KorAAAAYAAJ.

[8] Chenlin Gu. ***Lecture notes on metric space and Gromov-Hausdorff distance***. Sept. 2017. URL: https://chenlin-gu.github.io/notes/GromovHausdorff.pdf.

[9] Gady Kozma, Zvi Lotker, and Gideon Stupp. "The minimal spanning tree and the upper box dimension". In: ***Proceedings of the American Mathematical Society*** 134.4 (Sept. 2005), 1183–1187. ISSN: 1088-6826. DOI: 10.1090/s0002-9939-05-08061-5. URL: http://dx.doi.org/10.1090/S0002-9939-05-08061-5.

[10] Nick Kunz. ***fractal-dimensions***. https://github.com/NickMKunz/fractal-dimensions. 2024.

[11] Tom Leinster. ***The magnitude of metric spaces***. 2011. arXiv: 1012.5857 [math.MG].

[12] Tom Leinster and Mark W. Meckes. "The magnitude of a metric space: from category theory to geometric measure theory". In: ***Measure Theory in Non-Smooth Spaces***. De Gruyter Open, Dec. 2017, 156–193. ISBN: 9783110550832. DOI: 10.1515/9783110550832-005. URL: http://dx.doi.org/10.1515/9783110550832-005.

[13] R. Magnus. ***Metric Spaces: A Companion to Analysis***. Springer Undergraduate Mathematics Series. Springer International Publishing, 2022. ISBN: 9783030949464. URL: https://books.google.ch/books?id=E81kEAAAQBAJ.

[14] Clément Maria et al. "Rips complex". In: ***GUDHI User and Reference Manual***. GUDHI Editorial Board, 2016. URL: http://gudhi.gforge.inria.fr/doc/latest/group__rips__complex.html.

[15] Mark W. Meckes. "Positive definite metric spaces". In: ***Positivity*** 17.3 (Sept. 2012), 733–757. ISSN: 1572-9281. DOI: 10.1007/s11117-012-0202-8. URL: http://dx.doi.org/10.1007/s11117-012-0202-8.

[16] Mark W. Meckes. "Magnitude, Diversity, Capacities, and Dimensions of Metric Spaces". In: ***Potential Analysis*** 42.2 (Oct. 2014), 549–572. ISSN: 1572-929X. DOI: 10.1007/s11118-014-9444-3. URL: http://dx.doi.org/10.1007/s11118-014-9444-3.

[17] W. Rudin. ***Real and Complex Analysis***. Higher Mathematics Series. McGraw-Hill Education, 1987. ISBN: 9780070542341. URL: https://books.google.ch/books?id=Z_fuAAAAMAAJ.

[18] Benjamin Schweinhart. ***Persistent Homology and the Upper Box Dimension***. 2019. arXiv: 1802.00533 [math.MG].

[19] J. Michael Steele. "Growth Rates of Euclidean Minimal Spanning Trees with Power Weighted Edges". In: ***The Annals of Probability*** 16.4 (1988), pp. 1767 –1787. DOI: 10.1214/aop/1176991596. URL: https://doi.org/10.1214/aop/1176991596.

[20] Simon Willerton. "On the magnitude of spheres, surfaces and other homogeneous spaces". In: ***Geometriae Dedicata*** 168.1 (Feb. 2013), 291–310. ISSN: 1572-9168. DOI: 10.1007/s10711-013-9831-8. URL: http://dx.doi.org/10.1007/s10711-013-9831-8.

[21] Simon Willerton. ***Spread: a measure of the size of metric spaces***. 2015. arXiv: 1209.2300 [math.MG].

ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every written paper or thesis authored during the course of studies. In consultation with the supervisor, one of the following three options must be selected:

- [x] I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used no generative artificial intelligence technologies[1].
- [ ] I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used and cited generative artificial intelligence technologies[2].
- [ ] I confirm that I authored the work in question independently and in my own words, i.e. that no one helped me to author it. Suggestions from the supervisor regarding language and content are excepted. I used generative artificial intelligence technologies[3]. In consultation with the supervisor, I did not cite them.

Title of paper or thesis:

Fractal Dimensions in Theory and Practice

Authored by:
If the work was compiled in a group, the names of all authors are required.

Last name(s):	First name(s):
Kunz	Nick

With my signature I confirm the following:
- I have adhered to the rules set out in the Citation Guide.
- I have documented all methods, data and processes truthfully and fully.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for originality.

Place, date	Signature(s)
02.05.2024	N. K.

If the work was compiled in a group, the names of all authors are required. Through their signatures they vouch jointly for the entire content of the written work.

[1] E.g. ChatGPT, DALL E 2, Google Bard
[2] E.g. ChatGPT, DALL E 2, Google Bard
[3] E.g. ChatGPT, DALL E 2, Google Bard