Chaos game on fractals

Orsolya Kun

BSc Thesis

Supervisor: Dr. Balázs Bárány
BME-TTK
Institute of Mathematics
Department of Stochastics

2026

Contents

1 Introduction

Fractal geometry is an exciting field of mathematics. It describes complex shapes that traditional geometry cannot. Benoit Mandelbrot discovered that many objects in nature exhibit fractal properties [8]. These shapes are most commonly generated using Iterated Function Systems (IFS). An IFS consists of simple contraction mappings. In this thesis, we focus specifically on the Sierpinski gasket. It is a self-similar fractal, meaning that its magnified parts look exactly like the entire shape. These attractors can be rendered in several ways. One of the most efficient methods is the chaos game. This was introduced by Michael Barnsley [3]. The chaos game is a random iteration process. Here, we do not apply all mappings at every step. Instead, a function is chosen randomly based on a given probability distribution. Surprisingly, the random orbit of a single starting point traces out the entire fractal. Although the algorithm itself is very simple, its mathematical background is highly complex. Our main question is how quickly the game draws the fractal. How much time is required for the points to cover the attractor sufficiently densely? This rate of coverage is determined by the geometric properties of the fractal. However, investigating only the traditional dimensions of the set is not sufficient. The speed strongly depends on the invariant probability measure governing the process. Therefore, it is important to understand the dimension theory of measures. In particular, the Minkowski dimension of measures plays a key role in this. The goal of this thesis is to connect this abstract theory with practical computing. In our work, we build upon the recent results of Bárány, Jurga, and Kolossváry [1], as well as Falconer, Fraser, and Käenmäki [6], Jurga and Morris [10]. We examine how the probability vector affects the time required for fractal generation. We simulate the chaos game. In doing so, we also empirically test the theoretical estimates.

2 Theoretical background

This section presents the mathematical foundations of iterated function systems (IFS). We review the fundamental properties of contractions and self-similar attractors, introduce the symbolic space necessary to describe the dynamics, and finally, present the Hausdorff metric used to measure the convergence of sets.

2.1 Definition of Iterated function systems

Definition 2.1 (Hölder continuity). *Let (X, d_X) and (Y, d_Y) be metric spaces. A mapping $f : X \to Y$ is called α-Hölder continuous for some exponent $\alpha > 0$, if there exists a constant $C > 0$ such that for all $x, y \in X$:*

$$d_Y(f(x), f(y)) \leq C \cdot d_X(x, y)^\alpha. \tag{2.1}$$

Definition 2.2 (Lipschitz continuity). *Let (X, d) be a complete metric space. We call a function $f : X \to X$ Lipschitz continuous if there exists a positive constant $C > 0$ such that $d(f(x), f(y)) \leq C \cdot d(x, y)$ for all $x, y \in X$.*

Definition 2.3 (Contraction). *Let (X, d) be a complete metric space. We call a function $f : X \to X$ a contraction if there exists some $r \in [0, 1)$, called the contraction ratio, such that $d(f(x), f(y)) \leq r \cdot d(x, y)$ for all $x, y \in X$.*

Definition 2.4. *Let (X, d) be a complete metric space. An Iterated Function System (IFS) is defined as a finite collection of contraction mappings on X. Formally it is denoted as a set Φ, where $\Phi = \{f_i\}_{i \in I}$, such that I is finite and $f_i : X \to X$ is a contraction mapping with a contraction ratio $c_i < 1$.*

Theorem 2.1 (Hutchinson,[7]). *Let (X, d) be a complete metric space, and let $\Phi = \{f_i\}_{i \in I}$ be an Iterated Function System (IFS) consisting of contractive mappings. Then there exists a unique, non-empty compact set $\Lambda \subset X$, called the attractor of the IFS, such that:*

$$\Lambda = \bigcup_{i \in I} f_i(\Lambda). \tag{2.2}$$

Example 2.1 (Sierpiński gasket)**.** *The Sierpiński gasket (denoted by Λ) is the unique attractor generated by a system of three affine transformations $f_1, f_2, f_3 : \mathbb{R}^2 \to \mathbb{R}^2$, where*

$$f_1(x,y) = \left(\frac{x}{2}, \frac{y}{2}\right), \qquad f_2(x,y) = \left(\frac{x+1}{2}, \frac{y}{2}\right), \qquad f_3(x,y) = \left(\frac{x}{2} + \frac{1}{4}, \frac{y}{2} + \frac{\sqrt{3}}{4}\right).$$

The set Λ satisfies the fixed-point equation:

$$\Lambda = f_1(\Lambda) \cup f_2(\Lambda) \cup f_3(\Lambda). \tag{2.3}$$

Why do we model this using an IFS? This formulation stems directly from the geometric construction of the attractor Λ. At each step, the geometric shape is replaced by three half-sized copies of itself. This self-referential property is exactly what the Iterated Function System captures, and it perfectly explains the iterative generation process shown in Figure 1 below.

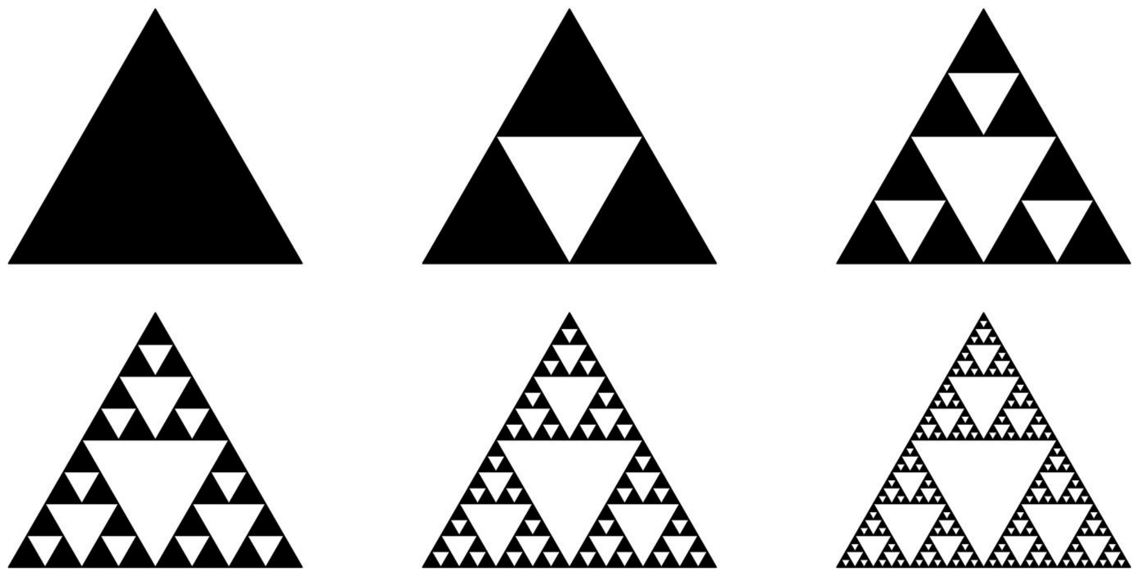

Figure 1: The firt 6 iterations of the Sierpinski gasket.

Definition 2.5. *Let Φ be an Iterated Function System (IFS), and let Λ be the attractor of Φ. If all the mappings $f \in \Phi$ are similarities — meaning there exists a constant $\lambda \in (0,1)$ for each mapping such that for all x, y in the metric space:*

$$d(f(x), f(y)) = \lambda d(x,y). \tag{2.4}$$

then we call the attractor Λ self-similar.

Definition 2.6 (Some classical fractal definition).

- *Middle-third Cantor set: It is generated on the real line* $\mathbb{R}$ *by an IFS consisting of two mappings:* $f_1(x) = \frac{1}{3}x$ *and* $f_2(x) = \frac{1}{3}x + \frac{2}{3}$. *Geometrically, this process recursively removes the open middle third of every line segment, starting from the interval* $[0, 1]$.

- *Sierpiński carpet: This attractor* $\Lambda \subset \mathbb{R}^2$ *is generated by 8 similarities, each with a contraction ratio* $\lambda = 1/3$. *If we consider the unit square* $[0, 1]^2$, *the IFS is defined as follows:*

$$f_{i,j}(x, y) = \left(\frac{x+i}{3}, \frac{y+j}{3} \right), \quad \text{where } (i, j) \in \{0, 1, 2\}^2 \setminus \{(1, 1)\}.$$

  This means we divide the square into 9 sub-squares and keep all of them except the central one (where $i = 1, j = 1$*).*

- *von Koch curve: It is generated by recursively replacing the middle third of a line segment with the two other sides of an outward-pointing equilateral triangle. Its IFS consists of* 4 *similarities in* $\mathbb{R}^2$, *all having a contraction ratio of* $1/3$, *which include scaling, translations, and rotations by* $60°$ *and* $-60°$.

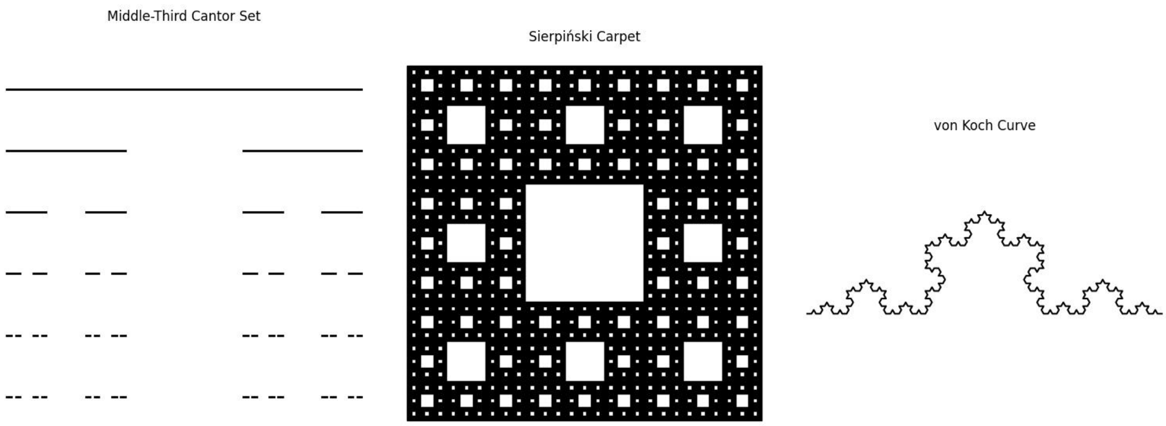

Figure 2: Visual representation of the mentioned classical fractals: the iterative steps of the Middle-Third Cantor Set (left), the self-similar structure of the Sierpiński Carpet (center), and the boundary line of the von Koch Curve (right).

2.2 The symbolic space

To analyze the dynamics of Iterated Function Systems, we introduce the symbolic space, which formally encodes the sequence of the applied transformations.

Definition 2.7. *Let $\Phi = \{f_1, \ldots, f_k\}$ be an IFS. The symbolic space Σ corresponding to the IFS is defined as the set of all infinite sequences over the alphabet $\{1, \ldots, k\}$:*

$$\Sigma := \{1, \ldots, k\}^{\mathbb{N}}. \tag{2.5}$$

The elements of Σ are denoted by $i = (i_1, i_2, \ldots)$ and $j = (j_1, j_2, \ldots)$. The subsets of Σ are the symbolic cylinders, which fix the first n coordinates of the sequences:

$$[i_1, \ldots, i_n] := \{j \in \Sigma : j_r = i_r \text{ for all } r \leq n\}. \tag{2.6}$$

To treat Σ as a geometric object, we equip it with a metric based on how long two sequences share the same prefix.

Definition 2.8. *For any two sequences $i, j \in \Sigma$, let $|i \wedge j| := \sup\{n \geq 0 : i_r = j_r \text{ for all } r \leq n\}$ denote the length of their longest common prefix. We define a metric ϱ on Σ as:*

$$\varrho(i, j) := k^{-|\mathbf{i} \wedge \mathbf{j}|} \tag{2.7}$$

with the convention that $k^{-\infty} = 0$ if $i = j$.

The most fundamental theoretical tool relating this abstract symbolic space to our physical geometric attractor $\Lambda \subset X$ is the natural projection map.

Definition 2.9. *Let $\Phi = \{f_1, \ldots, f_k\}$ be a strictly contracting IFS on a complete metric space X with attractor Λ. For any sequence $\mathbf{i} \in \Sigma$ and any arbitrary initial point $x_0 \in X$, the limit*

$$\pi(\mathbf{i}) := \lim_{n \to \infty} f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_n}(x_0) \tag{2.8}$$

exists and is entirely independent of the choice of the starting point x_0. The resulting mapping $\pi : \Sigma \to \Lambda$ is called the natural projection (or coding map).

2.3 Hausdorff metric

Definition 2.10. *The Hausdorff distance $d_H(A, B)$ between two non-empty sets A and B is defined as:*

$$d_H(A, B) := \inf\{\delta > 0 \mid A \subseteq B(B, \delta) \text{ and } B \subseteq B(A, \delta)\}, \tag{2.9}$$

where $B(A, \delta)$ denotes the δ-neighborhood of the set A, given by:

$$B(A, \delta) = \{y \in X \mid \exists x \in A,\, d(x, y) < \delta\}. \tag{2.10}$$

Alternatively, the Hausdorff distance can be expressed using the maximum of the directed distances:

$$d_H(A, B) = \max\{\delta_1, \delta_2\} \tag{2.11}$$

where $\delta_1 = \sup_{a\in A} \inf_{b\in B} d(a, b)$ and $\delta_2 = \sup_{b\in B} \inf_{a\in A} d(a, b)$.

Example 2.2. *Consider the set $A = [0, 1]^2$ in $\mathbb{R}^2$. The δ-neighborhood $B(A, \delta)$ with $\delta = \frac{1}{2}$ is formed by "inflating" the square by a radius of $\frac{1}{2}$ in all directions, resulting in a square with rounded corners.*

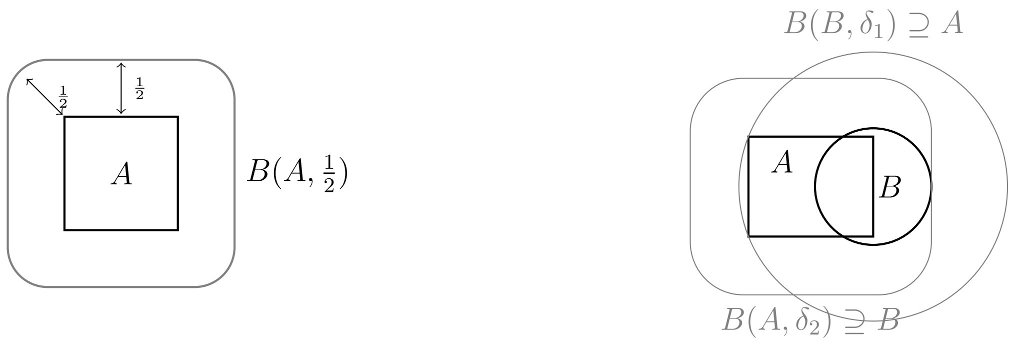

Figure 3: Left: The δ-neighborhood of a square $A = [0, 1]^2$. Right: Illustration of the Hausdorff metric between sets A and B, where $d_H(A, B) = \max\{\delta_1, \delta_2\}$.

Proposition 2.1 ([5])**.** *Let (X, d) be a metric space, and let $A, B, C \subseteq X$ be non-empty subsets. The Hausdorff distance satisfies the following properties:*

i) Symmetry: $d_H(A, B) = d_H(B, A)$.

ii) Triangle Inequality: $d_H(A, B) \leq d_H(A, C) + d_H(C, B)$.

These two properties establish that the Hausdorff distance is a premetric (or pseudometric) on the set of subsets of X.

Remark 2.1 (When does it become a true metric?)**.** *As established above, the Hausdorff distance is generally only a pseudometric because $d_H(A, B) = 0$ does not necessarily imply $A = B$. For instance, the distance between a set and its topological closure is zero. However, if we restrict our collection to the space of all non-empty compact subsets of a metric space, the identity of indiscernibles holds, and the Hausdorff distance becomes a true metric.*

To understand the topological structure of this space of compact sets, one of the most fundamental results is the Blaschke selection theorem.

Theorem 2.2 (Blaschke selection theorem,[5]). *Let $\mathcal{C}$ be an infinite collection of non-empty compact sets all lying in a bounded portion B of $\mathbb{R}^n$. Then there exists a sequence $\{E_j\}$ of distinct sets of $\mathcal{C}$ convergent in the Hausdorff metric to a non-empty compact set E.*

Proposition 2.2 ([7]). *Let (X, d) be a complete metric space, and let $f : X \to X$ be a contraction mapping with contraction ratio $\lambda \in (0, 1)$. Then for any non-empty subsets $A, C \subseteq X$, the following inequality holds for their Hausdorff distance:*

$$d_H(f(A), f(C)) \leq \lambda \cdot d_H(A, C). \tag{2.12}$$

Proposition 2.3 ([7]). *Let (X, d) be a complete metric space. For any finite collections of non-empty subsets $A_1, \ldots, A_n$ and $C_1, \ldots, C_n$ of X, the Hausdorff distance between their respective finite unions satisfies the following inequality:*

$$d_H\left(\bigcup_{i=1}^{n} A_i, \bigcup_{i=1}^{n} C_i\right) \leq \max_{i=1,\ldots,n} d_H(A_i, C_i). \tag{2.13}$$

3 The Concept of Fractal Dimension

Classical geometric measures (like standard length or area) are often inadequate for describing fractals. To see why, consider the two-dimensional Lebesgue measure (area), $Leb(\Lambda)$, of the Sierpiński gasket.

Example 3.1. *The Sierpiński gasket Λ is generated by an IFS of three similarities (Example 2.1). Each mapping scales the two-dimensional Lebesgue measure by a factor of $\lambda^2 = (1/2)^2 = 1/4$. Therefore, using the subadditivity of the measure on the fixed-point equation $\Lambda = \bigcup_{i=1}^{3} f_i(\Lambda)$, the total Lebesgue measure satisfies:*

$$Leb(\Lambda) = Leb\left(\bigcup_{i=1}^{3} f_i(\Lambda)\right) \leq \sum_{i=1}^{3} Leb(f_i(\Lambda)) = 3 \cdot \frac{1}{4} Leb(\Lambda). \tag{3.1}$$

Since $Leb(\Lambda) \leq \frac{3}{4} Leb(\Lambda)$ and the set Λ is bounded (thus its Lebesgue measure is finite, $Leb(\Lambda) < \infty$), this inequality can only hold if $Leb(\Lambda) = 0$.

This property holds more broadly: for any IFS in $\mathbb{R}^d$ with contraction ratios λ_i, if $\sum \lambda_i^d < 1$, then its d-dimensional Lebesgue measure is exactly $Leb_d(\Lambda) = 0$. To see why, we apply the subadditivity of the Lebesgue measure to the attractor's fixed-point property:

$$Leb_d(\Lambda) = Leb_d\left(\bigcup_{i=1}^{m} f_i(\Lambda)\right) \leq \sum_{i=1}^{m} Leb_d(f_i(\Lambda)) = \sum_{i=1}^{m} \lambda_i^d Leb_d(\Lambda) = \left(\sum_{i=1}^{m} \lambda_i^d\right) Leb_d(\Lambda).$$

Because the attractor Λ is a compact subset of $\mathbb{R}^d$, its measure is finite ($Leb_d(\Lambda) < \infty$). Given the assumption that $\sum \lambda_i^d < 1$, the inequality $Leb_d(\Lambda) \leq c \cdot Leb_d(\Lambda)$ with a constant $c < 1$ is mathematically only possible if $Leb_d(\Lambda) = 0$.

Because the Sierpiński gasket has a Lebesgue measure of zero but an infinite perimeter (at each construction step, the total boundary length is multiplied by $3 \times \frac{1}{2} = \frac{3}{2}$, thus diverging to infinity), its "true" geometric size cannot be meaningfully described by 1 or 2 dimensions. It must be measured in a fractional dimension strictly between 1 and 2.

3.1 Box-counting dimension

The box-counting dimension provides a formal way to measure how densely a set fills a metric space at shrinking scales.

Let (X, d) be a metric space, and let $E \subseteq X$ be a bounded subset. For any $\delta > 0$, let $N_\delta(E)$ denote the minimum number of open balls of radius δ required to cover the set E:

$$N_\delta(E) := \min\left\{ n \geq 0 \;\middle|\; \exists x_1, \ldots, x_n \in X : E \subseteq \bigcup_{i=1}^{n} B(x_i, \delta) \right\}. \tag{3.2}$$

Definition 3.1 (Box-counting dimension)**.** *The lower and upper box-counting dimensions of E are defined respectively as:*

$$\underline{\dim}_B(E) = \liminf_{\delta \to 0} \frac{\log N_\delta(E)}{-\log \delta} \tag{3.3}$$

$$\overline{\dim}_B(E) = \limsup_{\delta \to 0} \frac{\log N_\delta(E)}{-\log \delta} \tag{3.4}$$

If these two limits are equal, their common value is called the box-counting dimension of E, denoted by $\dim_B(E)$.

Remark 3.1. *A closed ball centred at $x \in X$ with radius $r > 0$ is denoted by $B(x, r)$.*

Remark 3.2. *The box-counting dimension is also referred to as the Minkowski dimension.*

Proposition 3.1 ([5])**.** *Let (X, d) be a metric space. The box-counting dimension satisfies the following fundamental properties:*

i) Monotonicity: For any subsets $A \subseteq C \subseteq X$, both the lower and upper box-counting dimensions preserve the subset relation:

$$\underline{\dim}_B(A) \leq \underline{\dim}_B(C) \quad \textit{and} \quad \overline{\dim}_B(A) \leq \overline{\dim}_B(C).$$

ii) Finite stability: The upper box dimension is finitely stable. For any finite collection of subsets $A_1, \ldots, A_n \subseteq X$:

$$\overline{\dim}_B\left(\bigcup_{i=1}^{n} A_i\right) = \max_{i=1,\ldots,n} \overline{\dim}_B(A_i).$$

iii) Relation to Lebesgue measure: If a set $A \subseteq \mathbb{R}^d$ has a positive d-dimensional Lebesgue measure ($Leb_d(A) > 0$), then its regular box-counting dimension exists and equals the ambient dimension:

$$\dim_B(A) = d.$$

Definition 3.2. *Let $\mathcal{B}$ be a countable collection of pairwise disjoint closed balls in a metric space X:*

- *It is called a δ-packing for $\delta > 0$ if all of the balls in $\mathcal{B}$ have radius δ.*
- *A δ-packing $\mathcal{B}$ is termed maximal if for every $x \in X$ there is a $B \in \mathcal{B}$ so that $B(x, \delta) \cap B \neq \emptyset$.*

Proposition 3.2 ([2])**.** *The box-counting dimension can be equivalently defined by replacing $N_\delta(E)$ with any of the following quantities. Let $E \subseteq \mathbb{R}^d$ be a bounded set:*

- *$N^1_\delta(E)$: The minimal number of axis-parallel cubes of side length δ needed to cover E.*
- *$N^2_\delta(E)$: The number of δ-mesh cubes (from a standard fixed grid of size δ) that intersect E.*
- *$N^3_\delta(E)$: The maximal number of pairwise disjoint balls of radius δ with centers in E.*

Substituting any of these $N^i_\delta(E)$ into the dimension formula yields the exact same limit, and thus the same box-counting dimension $\dim_B(E)$ (lower and upper respectively).

Remark 3.3. *Note that if $\mathcal{B}$ is a maximal δ-packing, then the collection with doubled radius, $2\mathcal{B} = \{B(x, 2\delta) : B(x, \delta) \in \mathcal{B}\}$, covers X.*

Let X be compact. The maximum cardinality of a δ-packing corresponds exactly to the quantity $N_\delta^3(X)$ introduced above:

$$N_\delta^3(X) = \max\{\#\mathcal{B} \mid \mathcal{B} \text{ is a } \delta\text{-packing}\} < \infty. \tag{3.5}$$

3.2 Hausdorff dimension

To define the Hausdorff dimension, we first need to construct the Hausdorff measure. Let (X, d) be a metric space. For any non-empty subset $U \subseteq X$, its diameter is defined as $|U| = \sup\{d(x, y) \mid x, y \in U\}$. If $U = \emptyset$, we define $|\emptyset| = 0$.

Definition 3.3 (Hausdorff measure)**.** *Let $E \subseteq X$, $s \geq 0$, and $\delta > 0$. We define the s-dimensional Hausdorff δ-approximation (or δ-content) of E as:*

$$H_\delta^s(E) = \inf\left\{\sum_{i\in I} |U_i|^s \;\middle|\; I \textit{ is countable}, \forall i \in I : |U_i| < \delta, \textit{ and } E \subseteq \bigcup_{i\in I} U_i\right\}. \tag{3.6}$$

The s-dimensional Hausdorff measure of E is then defined as the limit as the approximation scale δ approaches zero:

$$H^s(E) = \lim_{\delta\to 0} H_\delta^s(E) = \sup_{\delta>0} H_\delta^s(E). \tag{3.7}$$

Definition 3.4 (Metric outer measure)**.** *Let (X, d) be a metric space. An outer measure μ on X is called a metric outer measure if it satisfies the metric property:*

$$\mu(A \cup B) = \mu(A) + \mu(B) \tag{3.8}$$

for any sets $A, B \subseteq X$ such that their distance is strictly positive, i.e., $\inf\{d(a, b) \mid a \in A, b \in B\} > 0$.

Remark 3.4. *The significance of metric outer measures is established by Carathéodory's theorem. The theorem states that if an outer measure is metric, then every Borel set is measurable. This is one of the cornerstones of geometric measure theory. It is thanks to this theorem that we can reliably measure the most important sets (such as open, closed, and compact sets).*

Proposition 3.3 ([5])**.** *For any $s \geq 0$, the s-dimensional Hausdorff measure H^s is a metric outer measure on X.*

Corollary 3.1. *It follows directly from Carathéodory's theorem and Proposition 3.3 that all Borel subsets of X are H^s-measurable.*

Proposition 3.4 ([5])**.** *Let $E \subseteq X$ be a set and let $0 \leq s < t$. The Hausdorff measure satisfies the following properties:*

i) If $H^s(E) < \infty$, then $H^t(E) = 0$.

ii) If $H^t(E) > 0$, then $H^s(E) = \infty$.

Remark 3.5. *This property is essential because it ensures the existence of a critical threshold. As the parameter s increases, the measure $H^s(E)$ jumps from infinity to zero at exactly one point. This single transition point determines the value of the Hausdorff dimension. Without this property, the concept of dimension itself would not be well-defined.*

Definition 3.5 (Hausdorff dimension)**.** *Let $E \subseteq X$ be any set. The Hausdorff dimension of E, denoted by $\dim_H(E)$, is defined as the infimum of all values α for which the α-dimensional Hausdorff measure of E is zero:*

$$\dim_H(E) = \inf\{\alpha \geq 0 \mid H^\alpha(E) = 0\}. \tag{3.9}$$

Proposition 3.5 ([2])**.** *Let (X, d_X) and (Y, d_Y) be metric spaces. The Hausdorff dimension satisfies the following fundamental properties:*

i) Monotonicity: For any subsets $A \subseteq B \subseteq X$, the Hausdorff dimension preserves the subset relation:

$$\dim_H(A) \leq \dim_H(B).$$

This follows directly from the monotonicity of the Hausdorff measure, as $H^\alpha(A) \leq H^\alpha(B)$ for all $\alpha \geq 0$.

ii) Countable stability: Unlike the box-counting dimension, the Hausdorff dimension is countably stable. For any countable index set I and any collection of subsets $\{E_i\}_{i\in I}$ in X:

$$\dim_H\left(\bigcup_{i\in I} E_i\right) = \sup_{i\in I} \dim_H(E_i).$$

iii) Behavior under Hölder mappings: Suppose $f : X \to Y$ is an α-Hölder continuous mapping, then for any subset $A \subseteq X$:

$$\dim_H(f(A)) \leq \frac{\dim_H(A)}{\alpha}.$$

Proposition 3.6 ([2]). *Continuing from the previous properties, let E be a subset of a metric space.*

iv) Relation to Box-counting dimension: If E is a bounded set, its Hausdorff dimension is always bounded above by its lower box-counting dimension:

$$\dim_H(E) \leq \underline{\dim}_B(E).$$

Consequently, if the regular box-counting dimension exists, then $\dim_H(E) \leq \dim_B(E)$.

v) Relation to Lebesgue measure: For any subset $E \subseteq \mathbb{R}^d$, if its d-dimensional Lebesgue measure is strictly positive ($\text{Leb}_d(E) > 0$), then its Hausdorff dimension is equal to the ambient spatial dimension:

$$\dim_H(E) = d.$$

Lemma 3.1 ([5]). *Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : X \to Y$ be an α-Hölder continuous mapping. That is, there exist constants $c > 0$ and $\alpha > 0$ such that:*

$$d_Y(f(x), f(y)) \leq c \cdot d_X(x, y)^\alpha \quad \text{for all } x, y \in X.$$

Then, for any subset $A \subseteq X$ and any parameter $\beta \geq 0$, the following inequality holds for the Hausdorff measures:

$$H^\beta(f(A)) \leq c^\beta \cdot H^{\beta\alpha}(A). \tag{3.10}$$

Corollary 3.2. *Let (X, d_X) and (Y, d_Y) be metric spaces, and let $A \subseteq X$.*

1. *Lipschitz mapping: If $f : X \to Y$ is Lipschitz (which corresponds to an $\alpha = 1$-Hölder mapping), then it does not increase the Hausdorff dimension:*

$$\dim_H(f(A)) \leq \dim_H(A).$$

2. *Bi-Lipschitz mapping: If $f : X \to Y$ is bi-Lipschitz, then it exactly preserves the Hausdorff dimension:*

$$\dim_H(f(A)) = \dim_H(A).$$

3. *Similarity mapping: If $f : X \to Y$ is a similarity with ratio $\lambda > 0$, meaning $d_Y(f(x), f(y)) = \lambda \cdot d_X(x, y)$ for all $x, y \in X$, then it scales the α-dimensional Hausdorff measure by λ^α:*

$$H^\alpha(f(A)) = \lambda^\alpha H^\alpha(A).$$

Definition 3.6 (Similarity dimension)**.** *Let $\Phi = \{f_1, \dots, f_m\}$ be an IFS consisting of similarities on $\mathbb{R}^d$, where each mapping f_i has a contraction ratio $r_i \in (0, 1)$. The similarity dimension of the IFS, denoted by s, is defined as the unique positive real solution to the equation:*

$$\sum_{i=1}^{m} r_i^s = 1. \tag{3.11}$$

Lemma 3.2 ([2])**.** *Let Λ be the attractor of a self-similar IFS in $\mathbb{R}^d$. Then its Hausdorff dimension and upper box-counting dimension are bounded above by the minimum of the ambient dimension and the similarity dimension:*

$$\dim_H(\Lambda) \leq \overline{\dim}_B(\Lambda) \leq \min\{d, s\} \tag{3.12}$$

where s is the similarity dimension of the IFS. Moreover, its s-dimensional Hausdorff measure is finite, that is, $H^s(\Lambda) < \infty$.

Remark 3.6. *As we have seen, the similarity dimension s provides an upper bound for the Hausdorff dimension. However, if there is an overlap between the cylinder sets $f_i(\Lambda)$ of the attractor, the actual Hausdorff dimension can be strictly less than the similarity dimension. To ensure that $\dim_H(\Lambda) = s$, we need to restrict the amount of overlap between the pieces of the fractal.*

Definition 3.7 (Separation conditions)**.** *Let $\Phi = \{f_1, \dots, f_m\}$ be a contracting IFS in $\mathbb{R}^d$, and let Λ be its attractor. We define the following separation conditions:*

(a) Strong Separation Property (SSP): The IFS satisfies the SSP if the images of the attractor under the mappings are pairwise disjoint:

$$f_i(\Lambda) \cap f_j(\Lambda) = \emptyset \quad \textit{for all } i \neq j. \tag{3.13}$$

(b) Open Set Condition (OSC): The IFS satisfies the OSC if there exists a non-empty bounded open set $V \subset \mathbb{R}^d$ such that:

- *$f_i(V) \subseteq V$ for all $i = 1, \dots, m$;*
- *$f_i(V) \cap f_j(V) = \emptyset$ for all $i \neq j$.*

(c) Strong Open Set Condition (SOSC): The IFS satisfies the SOSC if the set V in the definition of the OSC can be chosen such that it intersects the attractor: $V \cap \Lambda \neq \emptyset$.

Remark 3.7. *The Strong Separation Property is a very restrictive assumption. For example, the middle-third Cantor set satisfies the SSP. However, many classical fractals, such as the Sierpiński gasket, the Sierpiński carpet, and the von Koch curve, do not satisfy the SSP because their cylinder sets intersect at their boundaries. Fortunately, these fractals do satisfy the Open Set Condition (OSC).*

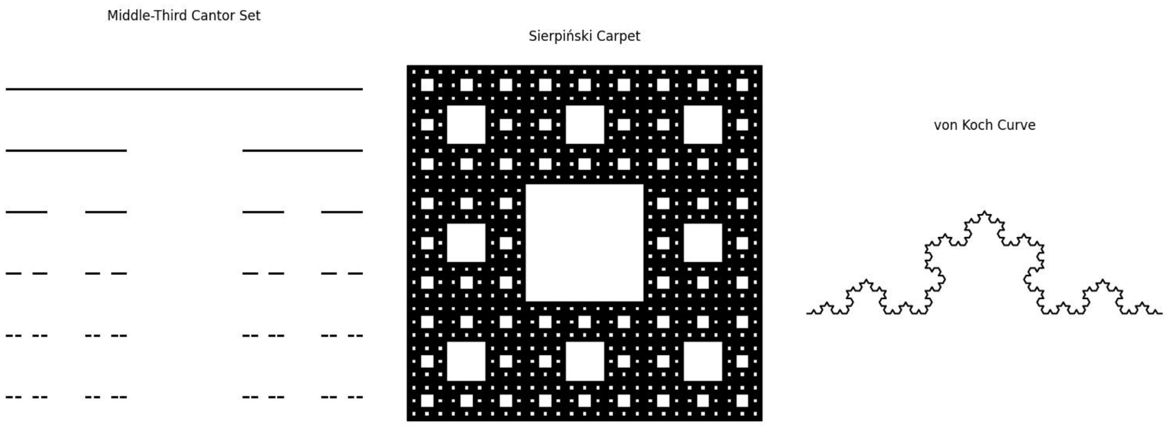

Figure 4: Visual representation of the mentioned classical fractals: the iterative steps of the Middle-Third Cantor Set (left), the self-similar structure of the Sierpiński Carpet (center), and the boundary line of the von Koch Curve (right).

Theorem 3.1 (Moran-Hutchinson,[9],[7])**.** *Assume that a self-similar Iterated Function System (IFS) $\Phi = \{f_1, \dots, f_m\}$ acts on $\mathbb{R}^d$ and satisfies the Open Set Condition (OSC). Let the contraction ratios of the similarities f_i be $0 < r_i < 1$ for $i = 1, \dots, m$. Let s be the similarity dimension:*

$$r_1^s + \cdots + r_m^s = 1.$$

Then, for the attractor Λ of the IFS Φ, the s-dimensional Hausdorff measure is strictly positive and finite:

$$0 < \mathcal{H}^s(\Lambda) < \infty. \tag{3.14}$$

Furthermore, the Hausdorff dimension and the box-counting dimension of the attractor coincide and are exactly equal to the similarity dimension s:

$$\dim_H(\Lambda) = \dim_B(\Lambda) = s. \tag{3.15}$$

In particular, this implies that the regular box-counting dimension exists.

Remark 3.8. *This theorem comes in very handy. It states that for fractals that satisfy the OSC, we do not need to calculate the notoriously difficult Hausdorff dimension from its definition using covers and infimums. Instead, we can simply solve the algebraic equation to find the similarity dimension s, which gives us the exact value for both the Hausdorff and the box-counting dimension.*

Example 3.2 (Dimension of the Sierpiński gasket)**.** *We have already established that the Sierpiński gasket Λ is generated by an IFS of three similarities ($m = 3$), each with a contraction ratio of $r_1 = r_2 = r_3 = 1/2$. We also verified that it satisfies the Open Set Condition. Therefore, applying the Moran-Hutchinson Theorem, its Hausdorff and box-counting dimensions are equal to the solution s of the equation:*

$$\begin{aligned} \left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^s + \left(\frac{1}{2}\right)^s &= 1 \\ 3 \cdot \left(\frac{1}{2}\right)^s &= 1 \\ 3 &= 2^s \\ s &= \frac{\log 3}{\log 2} \approx 1.58496 \end{aligned}$$

Thus, $\dim_H(\Lambda) = \dim_B(\Lambda) = \frac{\log 3}{\log 2}$.

3.3 Hausdorff dimension for measures and local dimension

While the Hausdorff dimension is primarily defined for sets, it is extremely useful to extend this concept to measures. This allows us to quantify how "spread out" or concentrated a probability measure is.

Definition 3.8 (Hausdorff dimensions of a measure)**.** *Let μ be a finite Borel measure on a metric space X.*

1. *The lower Hausdorff dimension of the measure μ, denoted by $\underline{\dim}_H(\mu)$, is defined as:*
$$\underline{\dim}_H(\mu) = \inf\{\dim_H(E) \mid E \textit{ is a Borel set and } \mu(E) > 0\}. \tag{3.16}$$

2. *The upper Hausdorff dimension of the measure μ, denoted by $\overline{\dim}_H(\mu)$, is defined as the infimum of the Hausdorff dimensions of all Borel sets E that have full measure:*
$$\overline{\dim}_H(\mu) = \inf\{\dim_H(E) \mid E \textit{ is a Borel set and } \mu(X \setminus E) = 0\}. \tag{3.17}$$

If the lower and upper Hausdorff dimensions coincide, their common value is simply referred to as the Hausdorff dimension of the measure, denoted by $\dim_H(\mu)$.

To connect the global dimension of a measure to its local behavior around specific points, we introduce the concept of local dimension (often called the pointwise dimension).

Definition 3.9 (Local dimension)**.** *Let μ be a finite Borel measure on X. For a point $x \in X$, the lower and upper local dimensions of μ at x are defined as:*

$$\underline{\dim}_{loc}(\mu, x) = \liminf_{r \to 0} \frac{\log \mu(B(x,r))}{\log r}, \tag{3.18}$$

$$\overline{\dim}_{loc}(\mu, x) = \limsup_{r \to 0} \frac{\log \mu(B(x,r))}{\log r}. \tag{3.19}$$

If these two limits are equal, the common value is called the local dimension of μ at x, denoted by $\dim_{loc}(\mu, x)$*.*

The following proposition establishes the fundamental connection between the local, point-by-point scaling behavior of a measure and its global Hausdorff dimension.

Proposition 3.7. *[2]*

Let μ be a finite Borel measure on a metric space X. The lower and upper Hausdorff dimensions of the measure satisfy:

$$\underline{\dim}_H(\mu) = \mu\text{-}ess\ inf\ \underline{\dim}_{loc}(\mu, x),$$

$$\overline{\dim}_H(\mu) = \mu\text{-}ess\ sup\ \underline{\dim}_{loc}(\mu, x).$$

Consequently, if the local dimension exists and equals a constant s for μ-almost every $x \in X$, then the Hausdorff dimension of the measure is exactly $\dim_H(\mu) = s$*.*

Theorem 3.2 (Converse of Frostman's lemma,[2])**.** *Let X be a compact metric space. The Hausdorff dimension of the set X can be fully characterized as the supremum of the Hausdorff dimensions of all positive Borel measures supported on it:*

$$\dim_H X = \sup \left\{ \underline{\dim}_H(\mu)) \mid \mu \text{ is a Borel measure with } \mu(X) > 0 \right\}. \tag{3.20}$$

3.4 Self-similar measures

Theorem 3.3 (Self-similar measure, [2]). *Let $\Phi = \{f_1, \dots, f_m\}$ be an IFS of contractive similarities on $\mathbb{R}^d$, and let $\mathbf{p} = (p_1, \dots, p_m)$ be a strictly positive probability vector (i.e., $p_i > 0$ and $\sum_{i=1}^m p_i = 1$). There exists a unique Borel probability measure $\mu_{\mathbf{p}}$, supported on the attractor Λ, satisfying the invariant equation:*

$$\mu_{\mathbf{p}} = \sum_{i=1}^{m} p_i \cdot \mu_{\mathbf{p}} \circ f_i^{-1}. \tag{3.21}$$

Such a measure $\mu_{\mathbf{p}}$ is called a self-similar measure.

Proposition 3.8 (Dimension of self-similar measures,[2]). *Assume that the IFS satisfies the Open Set Condition (OSC), and let $r_i \in (0,1)$ denote the contraction ratio of f_i. Then, for any probability vector $\mathbf{p}$, the local dimension of the self-similar measure $\mu_{\mathbf{p}}$ exists and is constant for $\mu_{\mathbf{p}}$-almost every point. Consequently, its Hausdorff dimension is given by the formula:*

$$\dim_H(\mu_{\mathbf{p}}) = \frac{\sum_{i=1}^k p_i \log p_i}{\sum_{i=1}^k p_i \log r_i}. \tag{3.22}$$

Example 3.3 (The natural measure of the Sierpiński gasket). *Recall that the Sierpiński gasket Λ is generated by an IFS of $m = 3$ similarities, each having a uniform contraction ratio of $c_i = 1/2$. Its similarity dimension is strictly $s = \frac{\log 3}{\log 2}$.*

For any self-similar IFS satisfying the Open Set Condition, the natural measure is defined as the self-similar measure $\mu_{\mathbf{p}}$ where the probability weights are chosen to be proportional to the contraction ratios raised to the power of the similarity dimension:

$$p_i = c_i^s.$$

For the Sierpiński gasket, substituting the values yields exactly:

$$p_1 = p_2 = p_3 = \left(\frac{1}{2}\right)^{\frac{\log 3}{\log 2}} = \frac{1}{3}.$$

This means the natural measure is generated by the uniform probability vector $\mathbf{p} = (1/3, 1/3, 1/3)$.

If we substitute these specific probabilities into the dimension formula from the previous proposition, we can calculate the dimension of this measure:

$$\dim_H(\mu_{\mathbf{p}}) = \frac{3 \cdot \left(\frac{1}{3}\log\left(\frac{1}{3}\right)\right)}{3 \cdot \left(\frac{1}{3}\log\left(\frac{1}{2}\right)\right)} = \frac{\log(1/3)}{\log(1/2)} = \frac{-\log 3}{-\log 2} = \frac{\log 3}{\log 2}.$$

Remarkably, the Hausdorff dimension of this natural measure exactly coincides with the Hausdorff dimension of the entire Sierpiński gasket Λ.

4 Minkowski dimension for measures

Before introducing the definition, it is useful to compare it to the well-established concept of the Hausdorff dimension of measures. It is a well-known fact that the Hausdorff dimension of a set can be approximated arbitrarily well from below by the Hausdorff dimensions of the measures supported on it. In the following, we will show that a similar principle applies to the Minkowski dimension: we will be able to approximate the Minkowski dimension of the set from above using measures.

4.1 Minkowski dimension for measures

The above definitions, and also the definitions of other set dimensions, extend naturally to all subsets of X by considering the restriction metric. Following the framework introduced by Falconer, Fraser, and Käenmäki [6], let μ be a fully supported Borel probability measure on X.

Definition 4.1. *We define the upper and lower Minkowski dimensions of the measure μ to be:*

$$\overline{\dim}_M(\mu) = \inf\{s \geq 0 \mid \exists c > 0 \textit{ such that } \mu(B(x,r)) \geq cr^s \textit{ for all } x \in X \textit{ and } 0 < r < 1\} \tag{4.1}$$

and

$$\begin{aligned}\underline{\dim}_M(\mu) = \inf\Big\{s \geq 0 \mid \exists c > 0 &\textit{ and a sequence } (r_n)_{n\in\mathbb{N}} \textit{ of positive real numbers} \\ &\textit{such that } \lim_{n\to\infty} r_n = 0 \textit{ and } \mu(B(x,r_n)) \geq cr_n^s \\ &\textit{for all } x \in X \textit{ and } n \in \mathbb{N}\Big\},\end{aligned} \tag{4.2}$$

respectively.

Remark 4.1. *The connection between the dimensions of a measure and the Minkowski dimension of its support X is somewhat delicate. As demonstrated by Falconer, Fraser, and Käenmäki, not requiring a uniform $0 < r < 1$ in the definition of the upper dimension would actually lead to the packing dimension instead of the Minkowski dimension.*

Proposition 4.1 ([6])**.** *Let (X, d) be a compact metric space and μ be a fully supported Borel probability measure on X. The upper and lower Minkowski dimensions of the measure μ can be equivalently expressed by the following limits:*

$$\overline{\dim}_M(\mu) = \limsup_{r\downarrow 0} \sup_{x\in X} \frac{\log \mu(B(x,r))}{\log r}, \tag{4.3}$$

$$\underline{\dim}_M(\mu) = \liminf_{r\downarrow 0} \sup_{x\in X} \frac{\log \mu(B(x,r))}{\log r}. \tag{4.4}$$

Proof. The proof follows from the definition of the upper and lower Minkowski dimensions. Let X be the support of the measure μ. We will prove the equivalence for the upper Minkowski dimension; the proof for the lower equality is analogous.

Let $s > \overline{\dim}_M(\mu)$ be arbitrary. By definition, there exists $c > 0$ such that $\mu(B(x,r)) \geq cr^s$ for all $x \in X$ and $0 < r < 1$. Taking logarithms and dividing by $\log r < 0$, we get for every $x \in X$:

$$\frac{\log \mu(B(x,r))}{\log r} \leq s + \frac{\log c}{\log r}$$

Taking the supremum over all $x \in X$ and then the limit superior as $r \downarrow 0$, the term $\frac{\log c}{\log r}$ vanishes, yielding:

$$\limsup_{r\downarrow 0} \sup_{x\in X} \frac{\log \mu(B(x,r))}{\log r} \leq s$$

Since $s > \overline{\dim}_M(\mu)$ was arbitrary, we conclude that the limit superior is at most $\overline{\dim}_M(\mu)$.

For the reverse inequality, we distinguish two cases based on the value of the upper Minkowski dimension.

Case 1: $\overline{\dim}_M(\mu) = 0$. Since μ is a probability measure, $\mu(B(x,r)) \leq 1$, meaning $\log \mu(B(x,r)) \leq 0$. Since $0 < r < 1$, we also know $\log r < 0$. Dividing a non-positive number by a negative number yields a non-negative number:

$$\frac{\log \mu(B(x,r))}{\log r} \geq 0$$

Taking the supremum over $x \in X$ and the limit superior as $r \downarrow 0$, the result is trivially ≥ 0. Thus, the inequality holds.

Case 2: $\overline{\dim}_M(\mu) > 0$. Let $0 \leq s < \overline{\dim}_M(\mu)$. By definition, for every $c > 0$ there exists a point $x_c \in X$ and a radius $0 < r_c < 1$ such that $\mu(B(x_c, r_c)) < c(r_c)^s$.

Taking logarithms and dividing by $\log r_c < 0$ gives:

$$\frac{\log \mu(B(x_c, r_c))}{\log r_c} > s + \frac{\log c}{\log r_c}$$

Because X is compact and μ is fully supported, the measure of any ball of a fixed radius $R > 0$ has a strictly positive minimum value, say $\inf_{x\in X} \mu(B(x, R)) = m_R > 0$. If r_c did not go to 0 as $c \to 0$, there would be some fixed radius $R > 0$ such that $r_c \geq R$ for arbitrarily small c. But this would mean $\mu(B(x_c, r_c)) \geq m_R > 0$, which directly contradicts $\mu(B(x_c, r_c)) < c(r_c)^s \to 0$ as $c \to 0$. Therefore, we must have $r_c \to 0$.

Because $r_c \to 0$ and $c \to 0$, the term $\frac{\log c}{\log r_c}$ is ≥ 0 (since $c < 1$ and $r_c < 1$, their logarithms are both negative). Thus:

$$\sup_{y\in X} \frac{\log \mu(B(y, r_c))}{\log r_c} \geq \frac{\log \mu(B(x_c, r_c))}{\log r_c} > s$$

Taking the limit superior as $r \downarrow 0$ establishes:

$$\limsup_{r\downarrow 0} \sup_{x\in X} \frac{\log \mu(B(x, r))}{\log r} \geq s$$

Since this holds for any $s < \overline{\dim}_M(\mu)$, the limit superior must be at least $\overline{\dim}_M(\mu)$. This establishes the equality for the upper Minkowski dimension. $\square$

Remark 4.2. *This characterization gives an easy way to compare the Minkowski dimensions to local dimensions of the measure, and therefore also to the Hausdorff and packing dimensions. Furthermore, a variational principle states that the Minkowski dimension of a set can be recovered entirely from the Minkowski dimensions of the measures supported on that set: specifically, the dimension of the set is exactly the infimum of the Minkowski dimensions of all fully supported measures.*

Theorem 4.1 (Variational principle for Minkowski dimension,[6])**.** *Let X be a compact metric space. The upper Minkowski dimension of the set X can be recovered entirely by finding the minimum dimension among all measures supported on it:*

$$\overline{\dim}_M(X) = \min\left\{\overline{\dim}_M(\mu) \;\middle|\; \mu \text{ is a fully supported finite Borel measure on } X\right\}. \quad (4.5)$$

An analogous claim holds for the lower Minkowski dimension $\underline{\dim}_M(X)$.

Proof. Let us first consider the claim for the upper Minkowski dimension. Let μ be a fully supported finite Borel measure on X and suppose $\overline{\dim}_M(\mu) < s < \infty$. It follows that there exists a constant $c > 0$ such that $\mu(B(x, r)) \geq cr^s$ for all $x \in X$ and $0 < r < 1$. If $\{B_i\}_{i=1}^N$ is an r-packing, then

$$Ncr^s \leq \sum_{i=1}^N \mu(B_i) \leq \mu(X).$$

Since this holds for every r-packing, we see that $N_r(X) \leq c^{-1}\mu(X)r^{-s}$ for all $0 < r < 1$ and hence, $\overline{\dim}_M(X) \leq s$. This proves one direction of the desired result.

To show the other direction, we may assume $\overline{\dim}_M(X) < \infty$ since otherwise there is nothing to prove. Let $k \in \mathbb{N}$ and choose a 2^{-k}-packing $\mathcal{B}_k$ such that $N_{2^{-k}}(X) = \#\mathcal{B}_k$. Note that, by the definition of $N_{2^{-k}}(X)$, $\mathcal{B}_k$ is maximal and hence, $2\mathcal{B}_k$ covers X. Write $N_k = N_{2^{-k}}(X)$ and $\{B(x_{k,i}, 2^{-k})\}_{i=1}^{N_k} = \mathcal{B}_k$. Fix $s > \overline{\dim}_M(X)$, choose $C \geq 1$ such that $N_k \leq Ck^{-2}2^{ks}$ for all $k \in \mathbb{N}$, and define

$$\mu = \sum_{k\in\mathbb{N}} k^{-2} \sum_{i=1}^{N_k} N_k^{-1}\delta_{x_{k,i}},$$

where δ_x is the Dirac measure at x. Since

$$\mu(X) = \sum_{k\in\mathbb{N}} k^{-2} \sum_{i=1}^{N_k} N_k^{-1} = \sum_{k\in\mathbb{N}} k^{-2} < \infty,$$

μ is a finite Borel measure on X. Furthermore, it is fully supported, as the following calculation will show. Given $x \in X$ and $0 < r < 1$, choose $k \in \mathbb{N}$ such that $2^{-k+1} < r \leq 2^{-k+2}$. Since $\{B(x_{k,i}, 2 \cdot 2^{-k})\}_{i=1}^{N_k}$ covers X, there exists $i \in \{1, \ldots, N_k\}$ such that $x_{k,i} \in B(x, 2 \cdot 2^{-k}) \subset B(x, r)$. Therefore,

$$\mu(B(x,r)) \geq k^{-2}N_k^{-1} \geq C^{-1}2^{-ks}.$$

Since $\mu(B(x,r)) > 0$ for every open ball, μ is indeed fully supported. The above inequality also proves $\overline{\dim}_M(\mu) \leq s$. Since $s > \overline{\dim}_M(X)$ was arbitrary, it follows that $\overline{\dim}_M(\mu) = \overline{\dim}_M(X)$, completing the proof for the upper dimension.

The claim for the lower Minkowski dimension is proved similarly. To see that $\underline{\dim}_M(X) \leq \underline{\dim}_M(\mu)$ for all fully supported finite Borel measures μ, just replace arbitrary radii $0 < r < 1$ by the appropriate sequence $(r_n)_{n\in\mathbb{N}}$ in the corresponding argument for the upper Minkowski dimension.

To see the other direction, let $(k_n)_{n\in\mathbb{N}}$ be a strictly increasing sequence of natural numbers such that $\underline{\dim}_M(X) = \lim_{n\to\infty} \log N_{2^{-k_n}}(X)/\log(2^{k_n})$. Let $n \in \mathbb{N}$ and choose a 2^{-k_n}-packing $\mathcal{B}_n$ such that $N_{2^{-k_n}}(X) = \#\mathcal{B}_n$. Note that, by the definition of $N_{2^{-k_n}}(X)$, $\mathcal{B}_n$ is maximal and hence $2\mathcal{B}_n$ covers X. Write $N_n = N_{2^{-k_n}}(X)$ and $\{B(x_{n,i}, 2^{-k_n})\}_{i=1}^{N_n} = \mathcal{B}_n$. Fix $s > \underline{\dim}_M(X)$, choose $C \geq 1$ such that $N_n \leq Ck_n^{-2}2^{k_n s}$ for all $n \in \mathbb{N}$, and define a finite Borel measure:

$$\mu = \sum_{n\in\mathbb{N}} k_n^{-2} \sum_{i=1}^{N_n} N_n^{-1}\delta_{x_{n,i}}.$$

Write $r_n = 2 \cdot 2^{-k_n}$ for all $n \in \mathbb{N}$ and notice that, for each $x \in X$ and $n \in \mathbb{N}$, we have

$$\mu(B(x, r_n)) \geq k_n^{-2} N_n^{-1} \geq C^{-1} 2^{-k_n s}.$$

Just as in the upper dimension case, this strictly positive measure ensures μ is fully supported. This also establishes $\underline{\dim}_M(\mu) \leq s$, and since s was arbitrary, $\underline{\dim}_M(\mu) = \underline{\dim}_M(X)$ as required. □

5 Chaos game

We can now define the main focus of this thesis, the chaos game. Let Λ be the attractor of an IFS $\Phi = \{f_1, \ldots, f_k\}$ acting on a complete metric space X. We choose a sequence of independently and identically distributed (i.i.d.) random variables $\omega_1, \omega_2, \ldots$ taking values in $\{1, \ldots, k\}$ with probability $\mathbb{P}(\omega_n = i) = p_i > 0$ where $\sum_{i=1}^{k} p_i = 1$.

We define the random orbit of a starting point $x \in X$ as the sequence x_n, where $x_0 = x$ and

$$x_n = f_{\omega_n}(x_{n-1}). \tag{5.1}$$

The random set $\mathcal{O}(x, \omega) = \{x_n : n \in \mathbb{N}\}$ is called the random orbit of x. The sequence of points $x_0, x_1, \ldots$ is generated by the chaos game.

Crucially, the chaos game defines a Markov process on the state space X. Let ν_p be the unique invariant Borel probability measure supported on Λ satisfying

$$\nu_p(A) = \sum_{i=1}^{k} p_i \nu_p(f_i^{-1}(A))$$

for any Borel set $A \subset X$. This invariant measure ν_p is exactly the stationary distribution for the chaos game. Consequently, the random orbit almost surely distributes according to this stationary measure as $n \to \infty$.

Definition 5.1. *A finite collection of contractions $\{f_0, \ldots, f_N\}$ on the complete metric space X along with a finite collection of probabilities $\{p_0, \ldots, p_N\}$ (where $p_i \geq 0$ and $\sum_{i=0}^{N} p_i = 1$) is called an IFS with probabilities.*

Remark 5.1. *To clarify the geometric meaning: a finite set of points $A = \{x_1, \ldots, x_n\}$ is called δ-dense in the target set F if for every point $y \in F$, there exists at least one generated point $x_j \in A$ such that the distance between them is strictly less than δ (i.e., $d(y, x_j) < \delta$).*

Theorem 5.1 ([4]). *Let X be a compact metric space and $\{f_i, p_i\}$ be a contractive IFSP. Choose some $x_0 \in X$, and define the sequence x_n by selecting $i \in \Sigma$ according to the probability measure $\mathbb{P}$ and setting*

$$x_n = f_{i_n} \circ \cdots \circ f_{i_1}(x_0).$$

Let μ be the invariant measure for the IFSP. Then we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} g(x_k) = \int_X g(x)\, d\mu(x)$$

for any continuous $g : X \to \mathbb{R}$ and for $\mathbb{P}$-almost all $\bar{\imath} \in \Sigma$.

Definition 5.2 (δ-covering time). *Let F be the attractor of an Iterated Function System consisting of the contractive maps $\{f_1, \dots, f_k\}$. Let v be a starting point, and let $i = (i_1, i_2, \dots)$ be an infinite sequence of chosen indices. We denote the position of the chaos game after n steps by:*

$$x_n = f_{i_n} \circ \cdots \circ f_{i_1}(v).$$

For a given spatial tolerance $\delta > 0$, the waiting time (or δ-covering time) $W_{\delta,v}(i)$ is defined as the minimum number of iterations required for the finite generated orbit to become δ-dense in the attractor F:

$$W_{\delta,v}(i) := \inf \left\{ n \geq 1 : \{f_{i_1} v, f_{i_2} f_{i_1} v, \dots, f_{i_n} \cdots f_{i_1} v\} \text{ is } \delta\text{-dense in } F \right\}.$$

Theorem 5.2 ([10]). *Let $(f_1, \dots, f_N)$ be an IFS of similitudes $S_i : \mathbb{R}^d \to \mathbb{R}^d$ with contraction ratios $r_i \in (0, 1)$ satisfying the OSC. Let s denote the similarity dimension of $(f_1, \dots, f_N)$, let $(p_1, \dots, p_N)$ be a nondegenerate probability vector, and define*

$$t := \max_{1 \leq i \leq N} \frac{\log p_i}{\log r_i}.$$

If the maximum in the definition of t is attained at a unique value $i \in \{1, \dots, N\}$, then there exists a constant $C > 0$ such that for every starting point $v \in F$ and every $0 < \delta < \min_{1 \leq i \leq N} r_i$,

$$C^{-1} \delta^{-t} \log\log\left(\frac{1}{\delta}\right) \leq \mathbb{E}(W_{\delta,v}) \leq C\, \delta^{-t} \left(\log\log\left(\frac{1}{\delta}\right)\right)^2.$$

If the maximum is not attained at a unique value of i, then there exists a constant $C > 0$ such that

$$C^{-1} \delta^{-t} \log\left(\frac{1}{\delta}\right) \leq \mathbb{E}(W_{\delta,v}) \leq C\, \delta^{-t} \log\left(\frac{1}{\delta}\right).$$

Remark 5.2. *The value t defined in the theorem is precisely the Minkowski dimension of the Bernoulli measure ν_p generated by the probabilities p_i and the similitudes f_i on the attractor. Morris and Jurga [10] showed that the expected time to cover the fractal up to a resolution δ grows roughly as δ^{-t}, where t corresponds to the expected time required to visit the "least accessible" part of the attractor. Later it was [1] generalized for a much wider class of IFS-s. They proved that the growth rate of this cover time is fundamentally determined by the Minkowski dimension of the measure driving the chaos game. For a self-similar IFS satisfying the OSC, the maximum local scaling exponent $t = \max_{1\leq i\leq N} \frac{\log p_i}{\log r_i}$ coincides exactly with the Minkowski dimension of the corresponding Bernoulli measure [1],[10].*

While the previous theorem describes the average behavior of the covering time through its expectation, it is equally important to understand how the algorithm behaves in individual random runs. The following theorem, established in [1], guarantees that the asymptotic behavior holds in a pointwise, almost sure sense for almost every generated sequence.

Theorem 5.3 ([1])**.** *Let $(f_1, \dots, f_N)$ be an IFS of similitudes satisfying the Open Set Condition, and let F be its attractor. Assume that the lower dimension of F is strictly positive. Let $\mathbf{p} = (p_1, \dots, p_N)$ be a nondegenerate probability vector, and let $\nu_{\mathbf{p}}$ be the corresponding self-similar measure. Then, for $\mathbb{P}$-almost every index sequence $i = (i_1, i_2, \dots) \in \Sigma$ and for every starting point $v \in F$, we have:*

$$\lim_{\delta\to 0} \frac{\log W_{\delta,v}(i)}{-\log\delta} = \dim_M(\nu_{\mathbf{p}})$$

where $\dim_M(\nu_{\mathbf{p}})$ denotes the Minkowski dimension of the push-forward measure $\nu_{\mathbf{p}}$.

Proof. We will illustrate the proof on the Sierpiński gasket however, it is available in a much more general context here [1]. Let the dynamical system be defined on the Sierpiński gasket, where the symbol space is $\{1, 2, 3\}^{\mathbb{N}}$ and the measure ν is uniform with $p_1 = p_2 = p_3 = \frac{1}{3}$. The definition of the hitting time is:

$$W_n(i, x_0) = \inf\{m \geq 0 : d_H(O_m(\mathbf{i}, x_0), \Lambda) \leq 2^{-n}\}.$$

Since the Sierpiński gasket is self-similar and satisfies the OSC, using the uniform distribution the Minkowski dimension is:

$$\dim_M(\nu_{\mathbf{p}}) = t := \max_{1\le i\le N} \frac{\log p_i}{\log r_i} = \frac{\log 3}{\log 2}.$$

Lower bound

We need to show that for almost every $\mathbf{i}$:

$$\nu\left(\left\{i \in \{1,2,3\}^{\mathbb{N}} : \liminf_{n\to\infty} \frac{\log W_n(i,x_0)}{n\log 2} < \frac{\log 3}{\log 2}\right\}\right) = 0.$$

This event can be written as the following union:

$$\bigcup_{k=1}^{\infty} \left\{i : \frac{\log W_n(\mathbf{i},x_0)}{n\log 2} < \frac{\log 3}{\log 2} - \frac{1}{k} \text{ for infinitely many } n\right\}.$$

By the Borel–Cantelli lemma, it is sufficient to show that for every $k \ge 1$:

$$\sum_{n=1}^{\infty} \nu\left(\left\{i : \frac{\log W_n(i,x_0)}{n\log 2} < \frac{\log 3}{\log 2} - \frac{1}{k}\right\}\right) < \infty.$$

Let us examine the probability of the n-th term. Rearranging the inequality:

$$\nu\left(\left\{i : W_n(i,x_0) < 3^n \cdot 2^{-\frac{n}{k}}\right\}\right).$$

This event is a subset of the event that the orbit hits a specific cylinder set of radius 2^{-n} within the given time (e.g., the neighborhood around the point $\pi(111\dots)$ consisting of all 1s):

$$\subseteq \left\{i : \exists 1 \le l \le 3^n \cdot 2^{-\frac{n}{k}} \text{ such that } f_{i_l} \circ \dots \circ f_{i_1}(x_0) \in B(\pi(111\dots), 2^{-n})\right\}.$$

We can estimate the probability from above using the union bound. The chance that the orbit falls into the cylinder set at a given time (i.e., the last n iterations are all 1s: $i'_l = 1, \dots, i'_{l-n} = 1$) is exactly 3^{-n}:

$$\le \sum_{l=1}^{\lfloor 3^n 2^{-n/k}\rfloor} \nu(i'l = 1, \dots, i'l - n = 1) = \lfloor 3^n 2^{-n/k}\rfloor \cdot 3^{-n} \le 3^n \cdot 2^{-\frac{n}{k}} \cdot 3^{-n} = 2^{-\frac{n}{k}}$$

Since $\sum_{n=1}^{\infty} 2^{-\frac{n}{k}} < \infty$, by the Borel–Cantelli lemma, almost surely:

$$\liminf_{n\to\infty} \frac{\log W_n(i,x_0)}{n\log 2} \geq \frac{\log 3}{\log 2}.$$

Upper bound

Now we need to show that:

$$\nu\left(\limsup_{n\to\infty} \frac{\log W_n(i,x_0)}{n\log 2} \leq \frac{\log 3}{\log 2}\right) = 1.$$

Which is equivalent to:

$$\nu\left(\limsup_{n\to\infty} \frac{\log W_n(i,x_0)}{n\log 2} > \frac{\log 3}{\log 2}\right) = 0.$$

$$\nu\left(\left\{\frac{\log W_n}{n\log 2} > \frac{\log 3}{\log 2} + \frac{1}{k} \text{ for infinitely many } n\right\}\right) = 0.$$

By the Borel-Cantelli lemma, it is enough to show that:

$$\sum_{n=1}^{\infty} \nu\left(\left\{\frac{\log W_n}{n\log 2} > \frac{\log 3}{\log 2} + \frac{1}{k}\right\}\right) < \infty.$$

Rearranging the condition we get $W_n > 3^n \cdot 2^{n/k}$. This means that the orbit does not yet form a 2^{-n}-cover of the attractor, so there exists at least one cylinder set $\bar{j} \in \{1,2,3\}^n$ of length n that we have not hit in the first $3^n \cdot 2^{n/k}$ steps:

$$\exists \bar{j} \in \{1,2,3\}^n : f_{i_l} \circ \cdots \circ f_{i_1}(x_0) \notin f_{j_1 \cdots j_n}(\Lambda) \text{ for all } 1 \leq l \leq 3^n 2^{n/k}.$$

Let us divide the time into independent blocks of length n. The number of blocks is at most $\frac{3^n . 2^{n/k}}{n}$. Applying the union bound to sum over all possible words $\bar{j} \in \{1,2,3\}^n$ of length n:

$$\leq \sum_{\bar{j}\in\{1,2,3\}^n} \nu\left(\bigcap_{l=1}^{\left\lfloor \frac{3^n 2^{n/k}}{n} \right\rfloor} \{[\tau'_{l\cdot n}, \ldots, \tau'_{(l-1)n+1}] \neq [j_1, \ldots, j_n]\}\right).$$

Since the blocks are independent, and the probability of a mismatch is $(1-3^{-n})$:

$$= \sum_{\bar{j}\in\{1,2,3\}^n} (1-3^{-n})^{\frac{3^n \cdot 2^{n/k}}{n}}.$$

Using the inequality $1 - x \le e^{-x}$:

$$= 3^n \cdot (1 - 3^{-n})^{\frac{3^n . 2^{n/k}}{n}} \le 3^n \cdot e^{-3^{-n} . \frac{3^n . 2^{n/k}}{n}} = 3^n \cdot e^{-\frac{2^{n/k}}{n}}.$$

$$a_n = 3^n \cdot e^{-\frac{2^{n/k}}{n}}$$

Applying the n-th root to the expression yields:

$$\sqrt[n]{a_n} = \left(3^n \cdot e^{-\frac{2^{n/k}}{n}}\right)^{\frac{1}{n}} = 3 \cdot e^{-\frac{2^{n/k}}{n^2}}$$

Since for any constant $k > 0$, the exponential function $(2^{n/k})$ grows asymptotically much faster than any polynomial term , the limit of the exponent is:

$$\lim_{n\to\infty} \frac{2^{n/k}}{n^2} = \infty$$

The limit of the entire expression is:

$$\lim_{n\to\infty} \sqrt[n]{a_n} = 3 \cdot e^{-\infty} = 3 \cdot 0 = 0$$

Since the sequence $3^n \cdot e^{-\frac{2^{n/k}}{n}}$ decreases as it is shown using Cauchy's root test , its sum is finite. Thus, by the Borel-Cantelli lemma, this event also occurs only finitely many times almost surely.

Conclusion
Combining the lower and upper bounds, for almost every sequence $\mathbf{i}$ and starting point x_0, it holds that:

$$\lim_{n\to\infty} \frac{\log W_n(i, x_0)}{n \log 2} = \frac{\log 3}{\log 2}.$$

□

Remark 5.3. *According to [1], the appearance of the Minkowski dimension of the measure,* $\dim_M(\nu_{\mathbf{p}})$, *in the almost sure convergence rate has a very intuitive geometric meaning. While the local dimension looks at the measure of a ball around a fixed point, the Minkowski dimension corresponds to the scale of the "least accessible part" of the attractor. In the chaos game, the algorithm is essentially bottlenecked by the regions of the fractal that receive the smallest probability mass. The covering time* $W_{\delta,v}$ *is dictated by how long we have to wait for the random orbit to finally visit these most improbable, heavily starved regions at resolution* δ.

5.1 Optimizing the Chaos Game

A natural question arises from the practical implementation of the chaos game: how should we choose the probability vector $\mathbf{p} = (p_1, \ldots, p_N)$ to generate the fractal as fast as possible?

Looking back at the expectation bounds, the growth rate of the expected waiting time is determined by the parameter $t = \max_{1 \leq i \leq N} \frac{\log p_i}{\log r_i}$. To minimize the waiting time, we must minimize this exponent t.

Proposition 5.1 ([2]). *Let s be the similarity dimension of the IFS. By the arithmetic-geometric mean inequality, we always have $t \geq s$. Equality holds if and only if the probabilities are chosen as:*

$$p_i = r_i^s \quad \textit{for all } i = 1, \ldots, N. \tag{5.2}$$

Remark 5.4. *This specific choice of probabilities, $p_i = r_i^s$, generates precisely the natural measure of the fractal, which we defined earlier. Therefore, the chaos game is statistically the fastest and most efficient when we play it using the natural measure. Any other choice of probabilities will result in some δ-balls in the attractor being substantially more difficult to access, slowing down the overall covering process.*

6 Practical implementation of the chaos game

6.1 The simulation setup

1. **Discretization of the fractal.** Since verifying the exact covering in a continuous metric space is analytically intractable, the program approximates the attractor (in this case, the Sierpinski gasket) with a sufficiently dense, finite set of points. Using the iterations of the given iterated function system (IFS), the algorithm generates a discrete reference grid whose resolution is significantly finer than the investigated radius δ. The primary theoretical condition for complete coverage is met when the Hausdorff distance between the random walk's generated orbit and the target attractor (represented by this fine grid) becomes less than or equal to δ.

2. **Execution of the random walk.** The chaos game is implemented as a Markov chain starting from the origin. In each iteration, the point applies a contracting transformation chosen according to the specified strict probabilities p_i. To continuously track the Hausdorff distance $d_H(O_n, \Lambda)$, the algorithm registers which points of the reference grid have been approximated within a distance of δ. The simulated walk terminates precisely when this Hausdorff metric condition is satisfied globally, meaning the unvisited reference points counter reaches zero.

3. **Data collection and expected value estimation.** To account for the high variance of the random process, the algorithm performs a large number (typically thousands) of independent simulations for a fixed δ, recording the total number of steps $T(\delta)$ required for the Hausdorff distance to drop below δ in each run. This aggregated empirical data forms the foundation for the subsequent analytical estimations.

6.2 Estimating the constant C

First let's recall Theorem 5.2 where, $\mathbb{E}(W_{\delta,v}) \leq C\,\delta^{-t}\log\left(\frac{1}{\delta}\right)$ we want to estimate constant C. Before estimating the constant C, we must consider the randomness of the chaos game. The number of steps required for a complete δ-cover, $T(\delta)$, is a random variable with high variance. Since the orbit is determined by independent random choices, the time required to satisfy the condition $d_H(O_n, \Lambda) \leq \delta$ varies significantly between simulations. This variability persists even with completely identical input parameters (r_i, p_i, δ). Due to chance, some orbits quickly reach even the most inaccessible parts of the fractal. Others, however, take a large number of redundant steps in already densely covered areas before covering the remaining δ-neighborhoods. Because of this high variance, a single simulation does not provide a reliable estimate for the expected cover time ($\mathbb{E}[T(\delta)]$). To determine the asymptotic behavior and the constant C, we must average the results of a large number of independent runs. The following figures aid in the geometric understanding of the process. These show the iterations under different probability vectors (p_i), illustrating how the gaps of the fractal are filled and highlighting the unpredictability of the exact time of coverage.

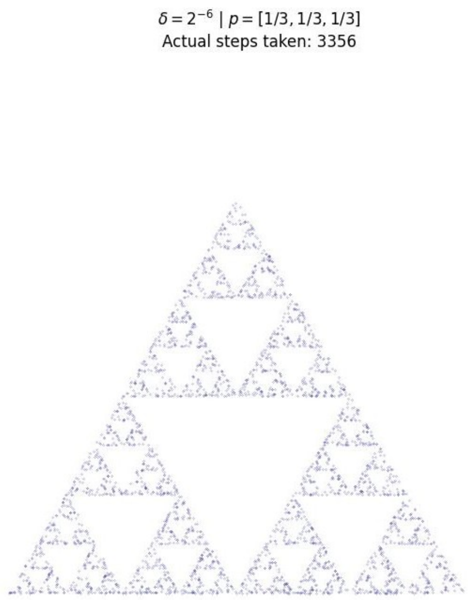

Figure 5: Chaos game simulated with uniform probabilities ($p_i = 1/3$). The δ-neighborhoods fill up relatively evenly across the attractor.

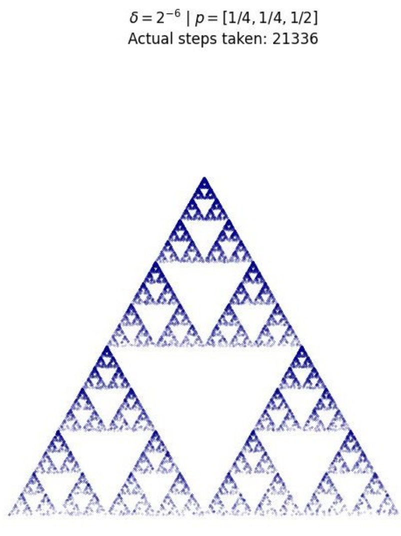

Figure 6: Chaos game simulated with non-uniform probabilities. The varying local dimensions cause certain regions to be covered significantly slower. This highly varying density geometrically illustrates why the covering time $T(\delta)$ can have such a large variance depending on the random orbit.

Following the theoretical framework, we now estimate the constant C empirically. To obtain reliable results, we ran exactly 1000 independent simulations for various (δ) and probability vectors. By decreasing δ, we observed the asymptotic behavior of the covering time as $\delta \to 0$. We conducted the analysis with both uniform ($p_i = 1/3$) and non-uniform probabilities. Based on the previous theorems, the asymptotic formula used to calculate C depends on the multiplicity of the maximum critical exponent $t_i = \frac{\log p_i}{\log r_i}$ (where $t = \max_i t_i$):"

The theoretical asymptotic formula used to calculate the constant C was determined by the multiplicity of the maximum among the critical exponents $t_i = \frac{\log p_i}{\log r_i}$ (where $t = \max_i t_i$), in accordance with the previously presented theorems:

- **Multiple maximums:** When the largest t_i value occurs at least twice, the formula $f(\delta) = \delta^{-t} \log(1/\delta)$ is applied as the theoretical base function. Crucially, both of our investigated configurations fall into this category. In the uniform distribution, all three exponents are equal (multiplicity of 3). In our specific non-uniform distribution ($p_1 = 1/4, p_2 = 1/4, p_3 = 2/4$), the maximum exponent $t = 2$ is shared by the first two transformations (multiplicity of 2).
- **Unique maximum:** For highly asymmetric probability distributions where the maximum of the critical exponents is strictly unique , the theoretical base function would take the form $f(\delta) = \delta^{-t} \log\log(1/\delta)$. While we acknowledge this theoretical case, our empirical study focuses on configurations belonging to the first category.

The estimated value of the constant C was determined as the ratio of the measured number of steps $T(\delta)$ to the theoretical baseline function $f(\delta) = \delta^{-t} \log(1/\delta)$. Due to the the right-skewness of the distribution and the distorting effect of extremely long random walk. We performed two types of estimation for each resolution δ:

$$C_{mean} \approx \frac{\text{Mean}(T(\delta))}{f(\delta)} \quad \text{and} \quad C_{median} \approx \frac{\text{Median}(T(\delta))}{f(\delta)}.$$

While C_{mean} strictly corresponds to the theoretical expected value required by the mathematical theorems, C_{median} is a more robust metric. By mitigating the impact of outliers, the latter describes the practical, typical coverage time of the chaos game.

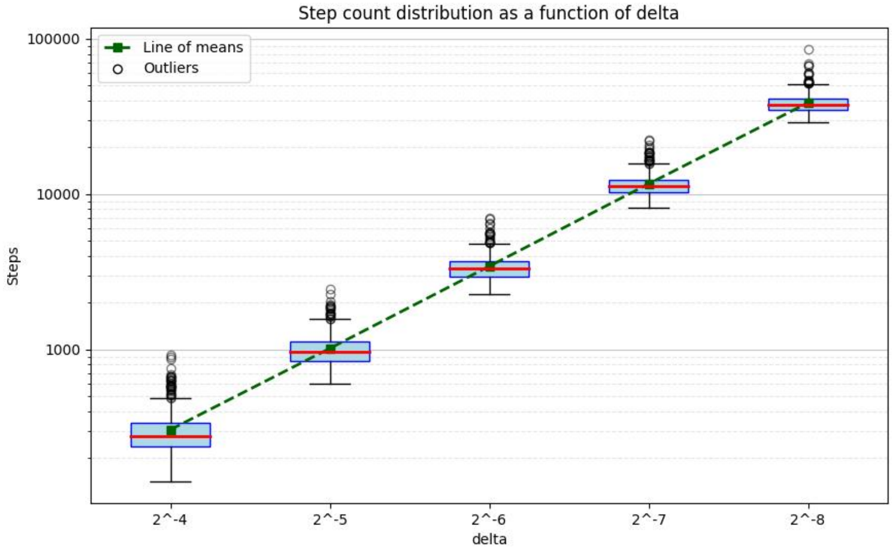

Figure 7: Box plot showing the distribution of simulation step counts as a function of delta.The upper black line represents the maximum, the lower one denotes the minimum, while the red line corresponds to the median.

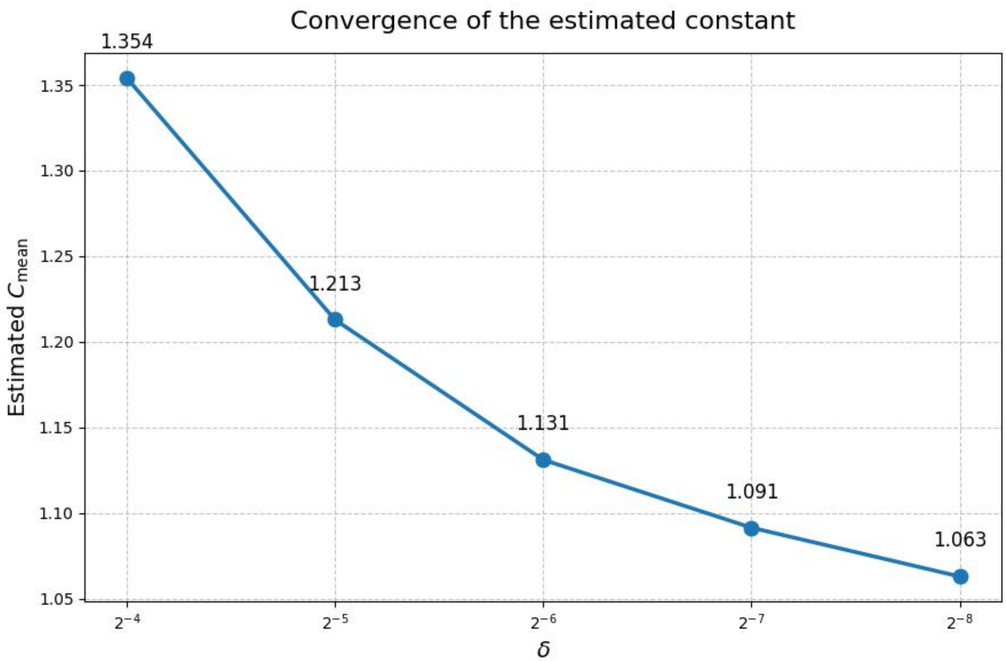

Figure 8: Convergence of the estimated empirical constant as δ decreases.

Although the mean is closely tied to the theoretical expected value, the distribution of coverage times is heavily right-skewed and contains many outliers (as shown by the box plot above). Extremely long random walks disproportionately distort the sample mean, obscuring the typical behavior of the chaos game. For this reason, we also estimated the constant C using the median:

$$C_{median} \approx \frac{\text{Median}(T(\delta))}{f(\delta)}.$$

The median is insensitive to outliers, thus providing a more reliable representation of typical simulations. Comparing C_{mean} and C_{median} reveals the distorting effect of extreme cases, and the difference between the two values serves as a strong indicator of the distribution's asymmetry as $\delta \to 0$.

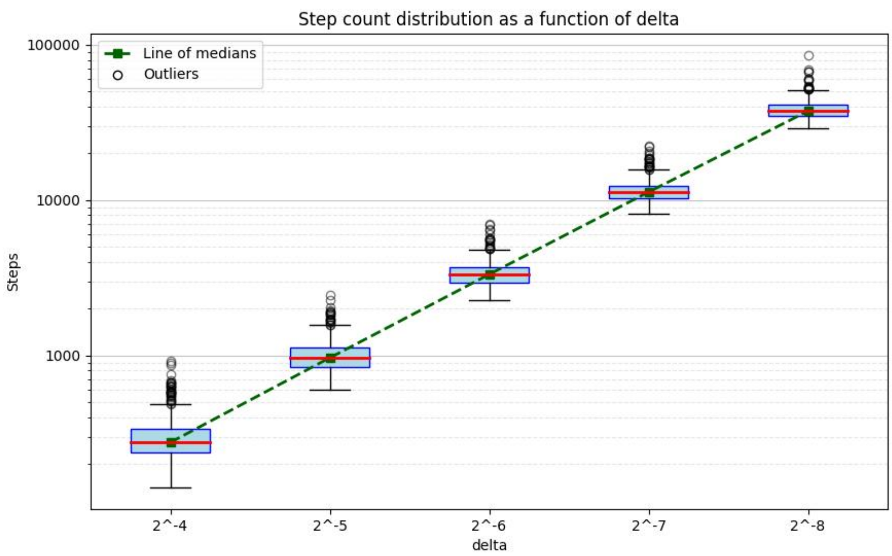

Figure 9: Box plot showing the distribution of simulation step counts as a function of delta (Unifrom distribution, median).The upper black line represents the maximum, the lower one denotes the minimum, while the red line corresponds to the median.

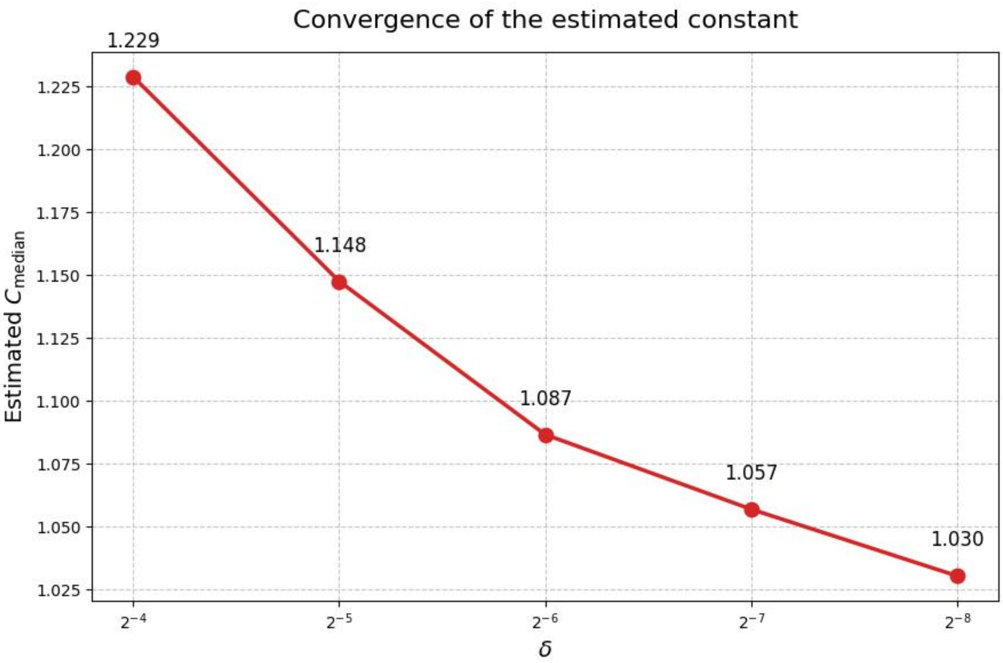

Figure 10: Robust estimation of the covering time constant based on the sample median. The use of the median provides a measure of central tendency that is resistant to the highly right-skewed nature of the covering time distribution.

6.3 Simulation results for the non-unifrom distribution

We now examine the case of a non-uniform probability vector $p = (1/4, 1/4, 1/2)$. In this case, the transformation with probability $p_3 = 1/2$ is chosen more frequently than the others. The critical exponents $t_i = \frac{\log p_i}{\log r_i}$ are as follows:

$$t_1 = t_2 = \frac{\log(1/4)}{\log(1/2)} = 2, \quad \text{and} \quad t_3 = \frac{\log(1/2)}{\log(1/2)} = 1.$$

The multiplicity of the maximum critical exponent ($t = \max_i t_i = 2$) is therefore 2. Since the maximum is not unique, in accordance with the preceding theorems, the baseline function describing the asymptotic behavior includes the logarithmic factor:

$$f(\delta) = \delta^{-t} \log(1/\delta).$$

To investigate the impact of the asymmetry and the convergence of the constant C, we repeated the simulation procedure: we executed 1000 independent runs for each value $\delta \in \{2^{-4}, \ldots, 2^{-8}\}$. Comparing this to the uniform case reveals whether the constant C is invariant to the choice of measure, or if the non-uniform distribution alters the C_{mean} and C_{median} estimates.

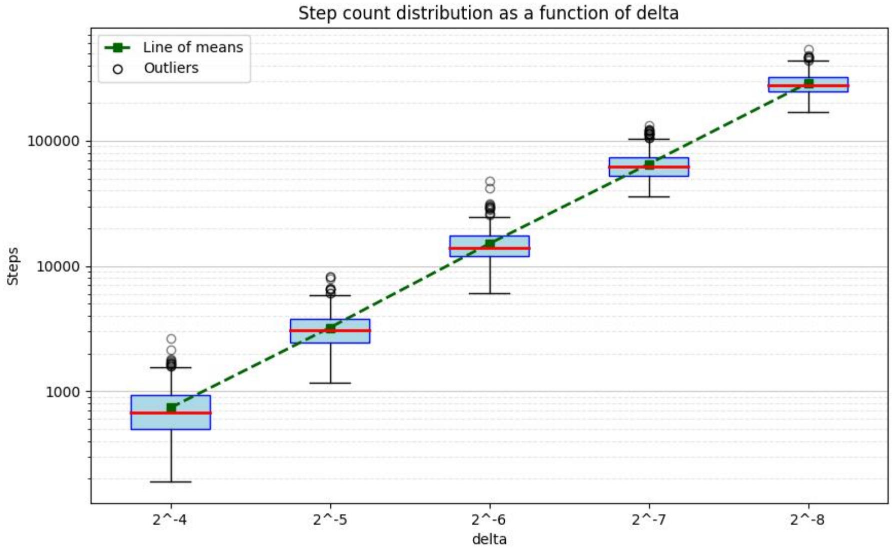

Figure 11: Box plot showing the distribution of simulation step counts as a function of delta (Non-uniform distribution, mean).The upper black line represents the maximum, the lower one denotes the minimum, while the red line corresponds to the median.

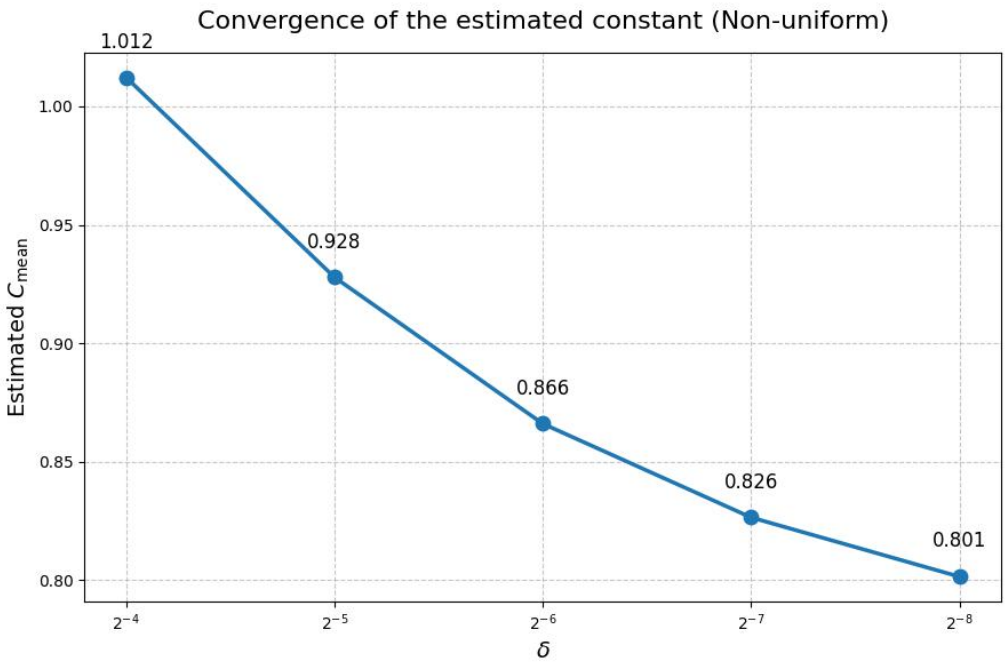

Figure 12: Empirical evaluation of the covering time constant based on the sample mean. While the sample mean rigorously connects to the theoretical expected value, its convergence behavior is visibly influenced by the right-skewed nature of the distribution and the geometric bottleneck created by the asymmetric probabilities.

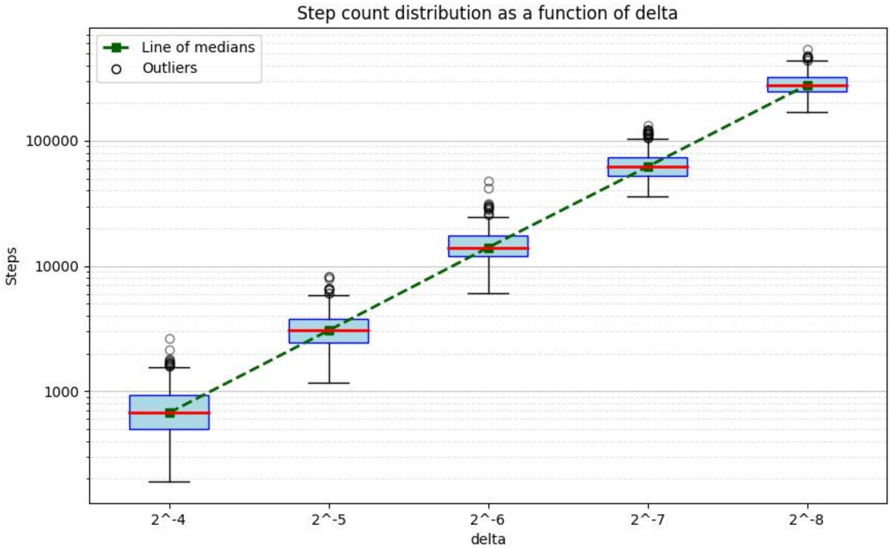

Figure 13: Box plot showing the distribution of simulation step counts as a function of delta (Non-uniform distribution, median).The upper black line represents the maximum, the lower one denotes the minimum, while the red line corresponds to the median.

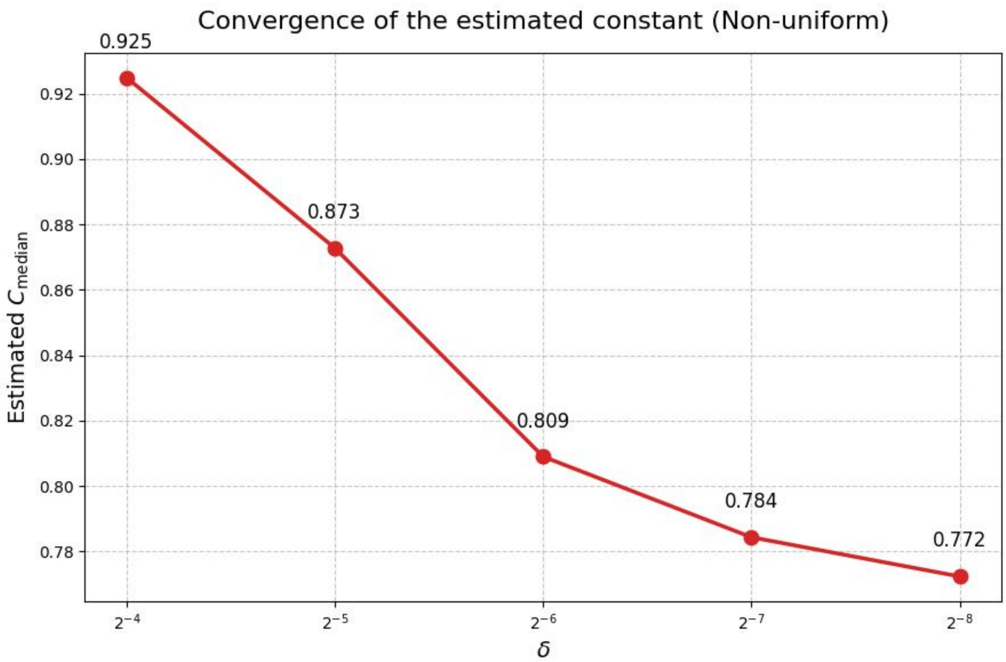

Figure 14: Robust empirical evaluation of the covering time constant based on the sample median. Utilizing the median effectively filters out the extreme outliers, which are heavily amplified by the geometric bottleneck of the asymmetric probabilities.

6.4 Comparison of the uniform and non-uniform cases

Comparing the simulation results highlights the dynamics of the chaos game. Although both probability vectors generate the Sierpiński gasket, the time required to achieve the δ-cover is not identical. This difference is demonstrated by the measured number of steps and the estimated constants C.

1. Magnitude of the covering time and critical exponents
The main difference between the two cases is the number of steps $T(\delta)$ required for the δ-cover. Examining the smallest delta ($\delta = 2^{-8}$), the median for the uniform distribution is approximately 40.000 steps. In contrast, for the non-uniform distribution, reaching the same δ-cover requires 200.000 steps.These simulation results support the theoretical calculations regarding the critical exponents. In the non-uniform case, due to the less frequently chosen transformations ($p_1 = p_2 = 1/4$), the random walk visits the corresponding regions less often. Consequently, the maximum critical exponent increases to $t = 2$ (compared to the uniform case's value of $t \approx 1.58$). According to mathematical theory, this higher exponent determines the total coverage time, which is confirmed by our empirical data.

2. Distribution asymmetry and the impact of outliers
Despite the differences in magnitude, the distribution of coverage times in both cases is right-skewed and contains many upper outliers. These extremely long coverage times occur when the orbit avoids certain parts of the fractal for an extended period, delaying the completion of the δ-cover. Due to this asymmetry, the estimates consistently yield $C_{mean} > C_{median}$. Although the mean is tied to the theoretical values, this difference confirms that, in practice, the median better characterizes the length of a typical simulation.

3. The Measure-dependence of the constant C
Finally, the convergence plots demonstrate that the asymptotic constant C is not a universal invariant of the fractal geometry itself, rather a parameter highly dependent on the applied probability measure.For example

- For the uniform measure, the median-based constant, at the smallest δ is $C \approx 1.030$.
- For the non-uniform measure, the median-based constant at the smallest δ is a lower value $C \approx 0.772$.

Our results confirm the theoretical expectations. The function $f(\delta)$ determines the scaling based on the critical exponents, while C reflects the specific measure. The asymmetric choice of probabilities increases the coverage time, and decreases the value of the constant as well.

References

[1] Balázs Bárány, Natalia Jurga, and István Kolossváry. On the convergence rate of the chaos game. *International Mathematics Research Notices*, 2023(5):4456–4500, 2023.

[2] Balázs Bárány, Károly Simon, and Boris Solomyak. *Self-similar and self-affine sets and measures*, volume 276 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2023.

[3] Michael F. Barnsley. *Fractals Everywhere*. Academic Press, 2nd edition, 1993.

[4] John H. Elton. An ergodic theorem for iterated maps. *Ergodic Theory and Dynamical Systems*, 7(4):481–488, 1987.

[5] Kenneth J. Falconer. *The Geometry of Fractal Sets*. Cambridge University Press, 1986.

[6] Kenneth J Falconer, Jonathan M Fraser, and Antti Käenmäki. Minkowski dimension for measures. *Proceedings of the American Mathematical Society*, 151(2):779–794, 2023.

[7] John E. Hutchinson. Fractals and self similarity. *Indiana University Mathematics Journal*, 30(5):713–747, 1981.

[8] Benoit B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman and Company, 1982.

[9] P. A. P. Moran. Additive functions of intervals and hausdorff measure. *Mathematical Proceedings of the Cambridge Philosophical Society*, 42(1):15–23, 1946.

[10] Ian D Morris and Natalia Jurga. How long is the chaos game? *Bulletin of the London Mathematical Society*, 53(6):1749–1765, 2021.