AN EXPLORATION OF FRACTAL DIMENSION

by

Dolav Cohen

B.S., California State University, Chico, 2010

A REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Mathematics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2013

Approved by:

Major Professor
Hrant Hakobyan

Abstract

When studying geometrical objects less regular than ordinary ones, fractal analysis becomes a valuable tool. Over the last 30 years, this small branch of mathematics has developed extensively. Fractals can be defined as those sets which have non-integral Hausdorff dimension. In this thesis, we take a look at some basic measure theory needed to introduce certain definitions of fractal dimensions, which can be used to measure a set's fractal degree. We introduce Minkowski dimension and Hausdorff dimension as well as explore some examples where they coincide. Then we look at the dimension of a measure and some very useful applications. We conclude with a well known result of Bedford and McMullen about the Hausdorff dimension of self-affine sets.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my major professor Dr. Hrant Hakobyan for his encouragement, patience and guidance during the preparation of this thesis. I also want to thank Dr. Nathan Albin and Dr. Diego Maldonado, who were on my committee, for all of their great support during my time in Kansas.

For all the love, support and inspiration from my family and friends through the years, I could not have done this without you. Thank you for everything.

Dedication

TO MY PARENTS

WITH LOVE

Chapter 1

INTRODUCTION

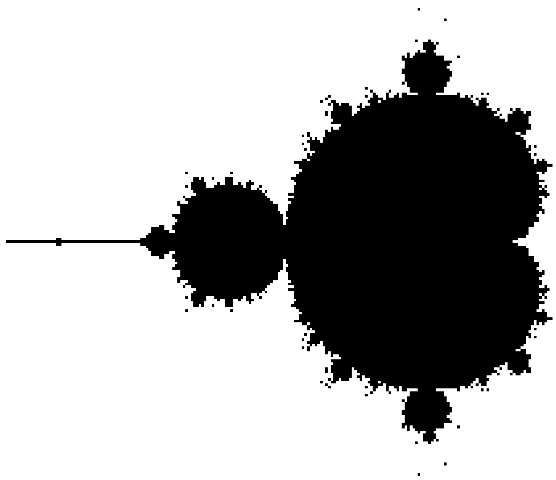
The basic idea of dimension that is used in our society is easily understood. Nonetheless, there is no single notion of dimension. Instead, there is a collection of alternative versions, each being applicable for different classes of mathematical spaces. Normally, a single mathematical object may have several distinctive notions of dimension that one can place on it. Dimensions are utilized as a tool to quantify the size of mathematical objects. For instance, in classical Euclidean geometry we can use the notion of dimension to visually perceive that a line segment is a one-dimensional object because it only has length. Yet a square is a two-dimensional object because it has both length and width and thus it has area. On the other hand, the space bounded by the sphere is three-dimensional because locally it has depth in addition to length and width.

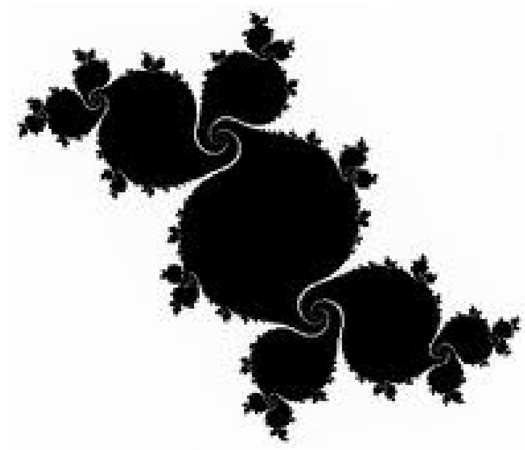
The development of fractional dimension has come a long way in the past century. More specifically, fractal sets have become much more important in different areas of sciences. According to Falconer[4], Mandelbrot recognized their use to model a wide variety of scientific phenomena from the molecular to the astronomical scales: the Brownian motion of particles, turbulence in fluids, the growth of plants, geographical coastlines and surfaces, the distribution of galaxies in the universe, and even fluctuations of price on the stock exchange. Dynamical systems and non-linear differential equations are additional areas of mathematics where sets of fractional dimension occur. An important characteristic of many mathematical spaces such as vector spaces and topological spaces is their dimension, which evaluates the complexity or degrees of freedom inherent in the space.

A fractal can be described as an object less regular than "ordinary" geometrical objects. The term 'fractal', derived from the Latin word *fractus*, meaning broken, was used as late as 1975 in the works of Benoit Mandelbrot, most notably in *The Fractal Geometry of Nature*[8]. He gave a conditional definition of a fractal as a set whose Hausdorff dimension is strictly greater than its topological dimension. However, he pointed out that the definition is undesirable as it eliminates particular highly irregular sets which clearly ought to be thought of as fractals. In this definition, the use of fractal dimension plays a big role and can be used to measure the "fractalness" of a fractal, thereby allowing comparisons between different fractals. Though the definition is relatively recent, examples of sets, now known as fractals in the sense of Mandelbrot, date back to the late 19th century. Weierstrass gave the first example of a nowhere differentiable function, the graph of which is a fractal. Even today however, the Hausdorff dimension of Weierstrass curves is not known[9]. The pictures of the Mandelbrot set and Julia sets from Dynamics were very important in the popularization of fractals in the 80's and 90's (see Fig. 1.1).

Fractal analysis is widely used today. Some examples include chaos theory and probability theory, among many others. Fractal objects and phenomena in nature such as mountains, coastlines and earthquakes are areas well studied by Mandelbrot. Much of the theory of Hausdorff dimension however, was developed by Abram S. Besicovitch and his school from the early 20th century, much before Mandelbrot.

In the theory of fractal dimensions and fractals there is still much to be explored. In fact, one of the hardest and most important problems in contemporary Harmonic Analysis is the well-known "Kakeya Problem"[5] : Is it true, that a set in $\mathbb{R}^N$ which contains a line segment in every direction has Hausdorff dimension equal to N? This is known to be true for sets in the plane $\mathbb{R}^2$. Such sets are known as Besicovitch sets, because Besicovitch was the first person who gave an example of a set in $\mathbb{R}^2$ which contained an interval in all directions





(b) The Julia Set

(a) The Mandelbrot Set

Figure 1.1: Two very well known fractals.

but had zero area. The "Kakeya Problem" is open even for Minkowski dimension.

In this thesis, we look at the two most common types of fractal dimensions known as Minkowski dimension and Hausdorff dimension. Then we introduce some of their important properties, as well as intriguing examples. One of the essential examples for both types of fractal dimensions is the middle thirds Cantor set. This example leads to the introduction of the Mass Distribution Principle, which helps calculate the lower bound for the Hausdorff dimension. Then we look at Billingsley's lemma, which is a finer version of the Mass Distribution Principle. We also define the dimension of a measure and take vertical slices of self-similar sets. Finally, we use the Law of Large Numbers, Billingsley's lemma and Boltzmann's Principle to prove the Bedford-McMullen formula for the Hausdorff dimension of self-affine sets.

Chapter 2

Minkowski Dimension

We begin by introducing one of the most popular notions of dimension in the study of fractals. Often referred to as the Minkowski dimension, this type of dimension has also been known as the box dimension, box-counting dimension, fractal dimension, metric dimension, capacity dimension or entropy dimension. It is commonly used and popular because it tends to be computationally easier to calculate or estimate.

The Minkowski dimension of a non-empty bounded subset of $\mathbb{R}^N$ is defined through an upper and lower dimension, which need not coincide. This type of dimension has many advantages; besides being easier to compute, the Minkowski dimension has a variety of equivalent definitions that may be used, which allows the analyst to choose whichever formulation is easier to work with for a given application. However, Minkowski dimension has some disadvantages. For instance, often the upper and lower dimensions are not equal and the Minkowski dimension of the set is not even well defined. Also, Minkowski dimension lacks certain desirable properties, such as countable stability.

2.1 Definitions

Below we consider subsets of Euclidean space $\mathbb{R}^N$ even though most of the definitions work in a general metric space X. The *distance* between points x and y in $\mathbb{R}^N$ will be denoted by $|x-y|$, i.e. if $x=(x_1,\ldots,x_N)$ and $y=(y_1\ldots,y_N)$ then

$$|x-y|=\sqrt{(x_1-y_1)^2+\ldots(x_N-y_N)^2}$$

Definition 2.1.1. An *open ball* and a *closed ball* with center $x\in\mathbb{R}^N$ and radius $r>0$ in $\mathbb{R}^N$ are defined, respectively as follows:

$$\begin{aligned} B(x,r) &= \{y:|y-x|<r\}, \\ \overline{B}(x,r) &= \{y:|y-x|\leq r\}. \end{aligned}$$

Note that in $\mathbb{R}^2$ a ball is a disc and in $\mathbb{R}^1$ a ball is an interval.

Definition 2.1.2. A *covering* of a set E is a finite or countable collection of open balls $E_1,E_2,E_3,\ldots$ such that

$$E\subset\bigcup_{i=1}^{\infty}E_i.$$

The balls are not necessarily all centered on a fixed point in E.

Definition 2.1.3. Given $\varepsilon>0$ and $E\subset\mathbb{R}^N$, let $N(E,\varepsilon)$ be the smallest number of $\varepsilon-balls$, i.e. balls of radius ε, needed to cover E. In other words, for $\varepsilon>0$

$$N(E,\varepsilon)=\min\{k\geq 1: \exists \text{ a finite covering } E_1,...,E_k \text{ of } E \text{ s.t. } |E_i|\geq\varepsilon, \text{ for } i=1,...k\},$$

where $|A|$ is the diameter of a set $A\subset E$. Recall, that the diameter of a set $E\subset\mathbb{R}^N$ is defined as follows

$$|E|=\mathrm{diam}(E)=\sup\{|x-y|:x,y\in E\}.$$

Definition 2.1.4. Suppose E is a bounded set in $\mathbb{R}^N$ and let $N(E,\varepsilon)$ be as before. We define the *upper Minkowski dimension* as

$$\overline{\dim_M}(E) = \limsup_{\varepsilon\to 0}\frac{\log N(E,\varepsilon)}{\log(1/\varepsilon)}, \tag{2.1.1}$$

and the *lower Minkowski dimension*

$$\underline{\dim}_M(E) \quad = \quad \liminf_{\varepsilon\to 0} \frac{\log N(E,\varepsilon)}{\log(1/\varepsilon)}. \tag{2.1.2}$$

If $\overline{\dim}_M(E) = \underline{\dim}_M(E)$ then we call it the *Minkowski dimension* of E, denoted $\dim_M(E)$.

Remark 2.1.5. According to Falconer[4], in this definition of the Minkowski dimension, the number $N(E,\varepsilon)$ can be replaced by any of the following numbers:

1. the smallest number of closed balls of radius ε needed to cover E;

2. the smallest number of cubes of side ε needed to cover E;

3. the smallest number of sets with diameter at most ε covering E;

4. the largest number of disjoint balls of radius ε with centers in E.

In addition, Falconer[6] mentions an alternative definition of the Minkowski dimension which seems to be quite different at first glance. This definition involves the n-dimensional volume of the ε-*neighborhood* E_ε of $E \subset \mathbb{R}^N$, given by

$$E_\varepsilon = \{x \in \mathbb{R}^N : |x-y| \le \varepsilon \text{ for some } y \in E\}$$

i.e. the set of points within distance ε of E. Then for $E \subset \mathbb{R}^N$ denote

$$\underline{\dim}_M(E) \quad = \quad n - \limsup_{r\to 0} \frac{\log \mathcal{L}^n(E_\varepsilon)}{\log \varepsilon}, \tag{2.1.3}$$

$$\overline{\dim}_M(E) \quad = \quad n - \liminf_{r\to 0} \frac{\log \mathcal{L}^n(E_\varepsilon)}{\log \varepsilon}, \tag{2.1.4}$$

and

$$\dim_M(E) = n - \lim_{r\to 0} \frac{\log \mathcal{L}^n(E_\varepsilon)}{\log \varepsilon}, \tag{2.1.5}$$

if this limit exists, where $\mathcal{L}^n$ is n-dimensional volume or n-dimensional Lebesgue measure.

2.2 Basic properties of Minkowski dimension

Here we state some of the main and well known properties of Minkowski dimension.

Lemma 2.2.1. *If $E, F, E_i \subset \mathbb{R}^N$ then the following properties hold*

1. ***Monotonicity**: If $E \subset F$ and $\dim_M(E)$ and $\dim_M(F)$ exist, then $\dim_M(E) \leq \dim_M(F)$.*

2. ***Finite stability**: For every $n \in \mathbb{N}$, $\overline{\dim_M}\bigcup_{i=1}^n E_i = \max_{i=1,\dots,n}\{\overline{\dim_M}(E_i)\}$.*

3. *Let $\overline{K}$ be the closure of K. Then*

$$\underline{\dim_M}(\overline{K}) = \underline{\dim_M}(K), \tag{2.2.1}$$

$$\overline{\dim_M}(\overline{K}) = \overline{\dim_M}(K). \tag{2.2.2}$$

We should note that the finite stability property does not hold for lower Minkowski dimension. Also, the extension of finite stability to countable stability is false even for the upper Minkowski dimension. This can be seen in the examples below.

Proof. All the properties follow quite easily from the definition of $N(E, \varepsilon)$.

1. Suppose $E \subset F$ then for every $\varepsilon > 0$ we have $N(E, \varepsilon) \leq N(F, \varepsilon)$. Hence, it follows by definition that
$$\lim_{\varepsilon\to 0} \frac{\log N(E,\varepsilon)}{\log(1/\varepsilon)} \leq \lim_{\varepsilon\to 0} \frac{\log N(F,\varepsilon)}{\log(1/\varepsilon)},$$
thus $\dim_M(E) \leq \dim_M(F)$.

2. Suppose we have $\overline{\dim_M}(E) := \limsup_{\varepsilon\to 0} \frac{\log N(E,\varepsilon)}{\log(1/\varepsilon)}$. Then from the following inequality
$$\max_{i=1,\dots,n}\{N(E_i, \varepsilon)\} \leq N\left(\bigcup_i E_i, \varepsilon\right) \leq n \max_{i=1,\dots,n}\{N(E_i, \varepsilon)\},$$

we have that

$$\limsup_{\varepsilon\to 0} \frac{\log N\left(\bigcup_i E_i, \varepsilon\right)}{\log(1/\varepsilon)} \le \limsup_{\varepsilon\to 0} n \frac{\log\{\max_{i=1,\dots,n}\{N(E_i,\varepsilon)\}\}}{\log(1/\varepsilon)}$$

Therefore, $\overline{\dim_M}\bigcup_i E_i = \max_{i=1,\dots,n}\{\overline{\dim_M}(E_i)\}$.

3. Let $\overline{B}_1, ..., \overline{B}_k$ be a finite collection of closed balls of radii ε. If the closed set

$$\bigcup_{i=1}^{k} \overline{B}_i = \overline{\bigcup_{i=1}^{k} B_i}$$

contains K, it also contains $\overline{K}$. Hence, the smallest number of closed balls of radius ε that cover K equals the smallest number required to cover the larger set $\overline{K}$. Therefore, (2.2.1) and (2.2.2) hold as desired.

□

In most of the examples below we will often find $N(E,\varepsilon)$ up to some errors. The next lemma shows that it will be enough for the calculation of Minkowski dimension.

We will say that two sequences A_n and B_n which approach $+\infty$ are "comparable" and write $A_n \approx B_n$ if there are constants $C_1 \ge 1$ and $0 < C_2 < \infty$ such that for every $n \in \mathbb{N}$ the following holds,

$$\frac{1}{C_1} B_n - C_2 \le A_n \le C_1 B_n + C_2.$$

Lemma 2.2.2. *If $A_n \approx B_n$ and ε_n is a sequence of positive numbers approaching 0 then*

$$\limsup_{n\to\infty} \frac{\log A_n}{\log(1/\varepsilon_n)} = \limsup_{n\to\infty} \frac{\log B_n}{\log(1/\varepsilon_n)} \tag{2.2.3}$$

$$\liminf_{n\to\infty} \frac{\log A_n}{\log(1/\varepsilon_n)} = \liminf_{n\to\infty} \frac{\log B_n}{\log(1/\varepsilon_n)}. \tag{2.2.4}$$

Proof. Since $A_n \le C_1 B_n (1 + \frac{C_2}{C_1 B_n})$, we will have

$$
\begin{aligned}
\limsup_{n\to\infty} \frac{\log A_n}{\log(1/\varepsilon_n)} &\leq \limsup_{n\to\infty} \frac{\log C_1 + \log B_n + \log(1+\frac{C_2}{C_1 B_n})}{\log(1/\varepsilon_n)} \\
&= 0 + \limsup_{n\to\infty} \frac{\log B_n}{\log(1/\varepsilon_n)} + 0,
\end{aligned}
$$

since C_1, C_2 are constants and B_n and $1/\varepsilon_n$ approach $+\infty$ as $n \to \infty$. Thus we obtain

$$
\limsup_{n\to\infty} \frac{\log A_n}{\log(1/\varepsilon_n)} \leq \limsup_{n\to\infty} \frac{\log B_n}{\log(1/\varepsilon_n)}.
$$

Similarly, since $A_n \geq \frac{B_n}{C_1}(1 - \frac{C_1 C_2}{B_n})$ we obtain

$$
\begin{aligned}
\limsup_{n\to\infty} \frac{\log A_n}{\log(1/\varepsilon_n)} &\geq \limsup_{n\to\infty} \frac{\log B_n - \log C_1 + \log(1-\frac{C_1 C_2}{B_n})}{\log(1/\varepsilon_n)} \\
&= \limsup_{n\to\infty} \frac{\log B_n}{\log(1/\varepsilon_n)} - 0 + 0.
\end{aligned}
$$

The two inequalities clearly imply (2.2.3), while (2.2.4) is proved in exactly the same way. □

2.3 Examples

In this section, we will calculate the Minkowski dimension of several different sets. All the calculations are based on the careful estimates of the numbers $N(E, \varepsilon)$. We begin with the simple example of the interval $[0, 1]$.

Example 2.3.1. $\dim_M([0, 1]) = 1$.

Proof. Suppose $X = [0, 1]$, then at least $\lfloor \frac{1}{\varepsilon} \rfloor$ intervals of length ε are needed to cover X. Then $\lfloor \frac{1}{\varepsilon} \rfloor + 2$ will do, i.e. $N(X, \varepsilon) \leq \frac{1}{\varepsilon} + 2$. Hence,

$$
\begin{aligned}
\overline{\dim}_M(X) &\leq \limsup_{\varepsilon\to 0} \frac{\log N(X,\varepsilon)}{\log(1/\varepsilon)} \leq \limsup_{\varepsilon\to 0} \frac{\log\left(\frac{1}{\varepsilon}+2\right)}{\log(1/\varepsilon)} \\
&= \limsup_{\varepsilon\to 0} \frac{\log(1+2\varepsilon) - \log\varepsilon}{-\log\varepsilon} = 1 - \limsup_{\varepsilon\to 0} \frac{\log(1+2\varepsilon)}{\log\varepsilon} \\
&\xrightarrow[\varepsilon\to 0]{} 1.
\end{aligned}
$$

Therefore, $\overline{\dim}_M(X) \leq 1$.

Similarly, using the fact that $N(X, \varepsilon) \geq \lfloor \frac{1}{\varepsilon} \rfloor - 2$, we obtain

$$\begin{aligned}
\underline{\dim}_M(X) &\geq \liminf_{\varepsilon \to 0} \frac{\log N(X,\varepsilon)}{\log(1/\varepsilon)} \geq \liminf_{\varepsilon \to 0} \frac{\log\left(\frac{1}{\varepsilon} - 2\right)}{\log(1/\varepsilon)} \\
&= \liminf_{\varepsilon \to 0} \frac{\log(1-2\varepsilon) - \log \varepsilon}{-\log \varepsilon} = 1 - \liminf_{\varepsilon \to 0} \frac{\log(1-2\varepsilon)}{\log \varepsilon} \\
&\underset{\varepsilon \to 0}{\longrightarrow} 1.
\end{aligned}$$

Therefore, $\underline{\dim}_M(X) \geq 1$ and the two inequalities give the desired result, that is,

$$\dim_M([0,1]) = 1.$$

$\square$

2.3.1 A countable set of positive dimension

Now, in light of the equalities (2.2.1) and (2.2.2) it may not be surprising to have examples of countable sets with positive dimension. Indeed, if $E = \mathbb{Q} \cap [0,1]$ is the set of rationals in the unit interval then $\dim_M(E) = \dim_M(\overline{E}) = \dim_M([0,1]) = 1$.

The next example illustrates the failure of countable sub-additivity for the Minkowski dimension, perhaps in a striking way. A seemingly very small set, which coincides with its own closure, has a large (positive) Minkowski dimension.

Lemma 2.3.2. *Let* $E = \{0\} \bigcup \{1, 1/2, 1/3, ..., 1/n, ...\}$. *Then* $\dim_M(E) = \frac{1}{2}$.

Proof. Suppose $\varepsilon \in (0,1)$. To estimate $N(E, \varepsilon)$ choose $n \in \mathbb{N}$ so that

$$x_n - x_{n+1} \leq \varepsilon < x_{n-1} - x_n.$$

Since

$$x_n - x_{n+1} = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$$

is a monotone decreasing sequence, this is possible and we have the estimates

$$\frac{1}{(n+1)^2} < \frac{1}{n(n+1)} \leq \varepsilon < \frac{1}{(n-1)n} < \frac{1}{(n-1)^2}.$$

To cover the set E we will cover the interval $(0, x_n)$ by intervals of length ε and will need $n-1$ more intervals to cover the rest of E. The number of intervals of length ε needed to cover $(0, x_n)$ is less than

$$\frac{x_n}{\varepsilon} = \frac{\frac{1}{n}}{\varepsilon} \leq \frac{\frac{1}{n}}{\frac{1}{n(n+1)}} = n+1.$$

Thus

$$N(E,\varepsilon) \leq (n-1) + (n+1) = 2n = \frac{2n}{n-1} \cdot (n-1) \leq \frac{2n}{n-1} \left(\frac{1}{\varepsilon}\right)^{1/2}.$$

Hence,

$$\begin{aligned}
\limsup_{\varepsilon \to 0} \frac{\log N(E,\varepsilon)}{\log(1/\varepsilon)} &\leq \limsup_{\varepsilon \to 0} \frac{\log\left(\frac{2n}{n-1}\right)\left(\frac{1}{\varepsilon}\right)^{1/2}}{\log(1/\varepsilon)} \\
&= \limsup_{\varepsilon \to 0} \frac{\log\left(\frac{2n}{n-1}\right) + \frac{1}{2}\log\frac{1}{\varepsilon}}{\log(1/\varepsilon)} = \limsup_{\varepsilon \to 0} \frac{\log\left(\frac{2n}{n-1}\right)}{\log(1/\varepsilon)} + \frac{1}{2}
\end{aligned}$$

Now, since $n \approx (\sqrt{\varepsilon})^{-1}$ for small ε, we have that $n \to \infty$ if and only if $\varepsilon \to 0$ and thus

$$\limsup_{\varepsilon \to 0} \frac{\log\left(\frac{2n}{n-1}\right)}{\log(1/\varepsilon)} = 0.$$

And we obtain that

$$\overline{\dim}_M(E) = \limsup_{\varepsilon \to 0} \frac{\log N(E,\varepsilon)}{\log(1/\varepsilon)} \leq \frac{1}{2}.$$

To obtain the opposite inequality, we note that we will need at least $n-1$ balls of length ε to cover E, i.e.

$$N(E,\varepsilon) \geq n-1 \geq \frac{n-1}{n+1}\left(\frac{1}{\varepsilon}\right)^{1/2}.$$

Hence,

$$\begin{aligned}\liminf_{\varepsilon\to 0}\frac{\log N(E,\varepsilon)}{\log(1/\varepsilon)} &\geq \liminf_{\varepsilon\to 0}\frac{\log(n-1)}{\log(1/\varepsilon)} \geq \liminf_{\varepsilon\to 0}\frac{\log\left(\frac{n-1}{n+1}\right)\left(\frac{1}{\varepsilon}\right)^{1/2}}{\log(1/\varepsilon)} \\ &= \liminf_{\varepsilon\to 0}\frac{\log\left(\frac{n-1}{n+1}\right)+\frac{1}{2}\log\left(\frac{1}{\varepsilon}\right)}{\log(1/\varepsilon)} = \frac{1}{2}\end{aligned}$$

The last inequality above is true since $n \to \infty$ as before and we get that $\log\left(\frac{n-1}{n+1}\right) \to 0$ as $\varepsilon \to 0$. Therefore, $\underline{\dim}_M(E) \geq \frac{1}{2}$ and thus, $\dim_M(E) = \frac{1}{2}$. □

2.3.2 Middle thirds Cantor set

The following is one of the most common examples used for calculating fractal dimension. It turns out, that for this next example the Minkowski dimension agrees with the Hausdorff dimension, which happens to be the case for most self-similar sets[4]. First we will construct the set, which is known as the middle thirds Cantor set, the triadic Cantor set or simply the Cantor ternary set, then we will calculate its Minkowski dimension.

Example 2.3.3. (The middle thirds Cantor Set): We begin with the closed interval $[0,1]$ which we will denote by $I_{0,1}$ and divide it into three subintervals of equal length. Remove the middle open interval of length $\frac{1}{3}$, centered at $\frac{1}{2}$ and call it $J_{1,1}$. Denote the components of the remaining set $I_{1,1}$ and $I_{1,2}$ such that

$$I_{1,1} = \left[0, \frac{1}{3}\right] \text{ and } I_{1,2} = \left[\frac{2}{3}, 1\right].$$

Next, from the middle of $I_{1,i}$ remove the open interval $J_{2,i}$ of length $\frac{1}{3}$, for $i = 1, 2$. Then

$$I_{2,1} = \left[0, \frac{1}{3^2}\right], I_{2,2} = \left[\frac{2}{3^2}, \frac{3}{3^2}\right], I_{2,3} = \left[\frac{6}{3^2}, \frac{7}{3^2}\right], \text{ and } I_{2,4} = \left[\frac{8}{3^2}, 1\right].$$

Continue in this same fashion. Let

$$\mathcal{C} = \bigcap_{k=0}^{\infty}\bigcup_{j=1}^{2^{k-1}} I_{k,j}$$

whence

$$I_{k,j} = I_{k+1,2j-1} \bigcup J_{k+1,j} \bigcup I_{k+1,2j},$$

$$|J_{k,j}| = \frac{1}{3}|I_{k-1,j}| \text{ and } |I_{k+1,j}| = |I_{k+1,j'}| \ \forall j, j',$$

i.e. the length of the removed interval is $\frac{1}{3}$ of $|I_{k-1,j}|$ and the length of any of the remaining intervals, $|I_{k+1,j}|$, is the same as $|I_{k+1,j'}|$, for any j and j'. The first four iterations are shown in Fig. 2.1.

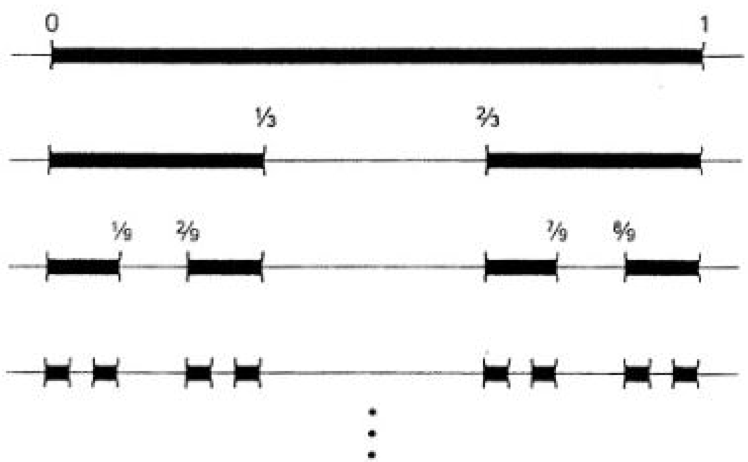


Figure 2.1: The middle thirds Cantor set

Lemma 2.3.4. *The middle thirds Cantor set $\mathcal{C}$ is uncountable.*

Proof. Consider the second generation $I_{1,j}$ for $j = 1, 2$ of $\mathcal{C}$ and the binary sequences $(x_i)_{i=1}^{\infty}$ with $x_i \in \{0, 1\}$. For every $c \in \mathcal{C}$ we let $x_1 = 0$ if c belongs to the left segment of $I_{1,j}$ for $j = 1, 2$ and $x_1 = 1$ if c is found in the right segment of $I_{1,j}$ for $j = 1, 2$. Next, we need to consider in which of the two possible segments of $I_{2,j}$ for $j = 1, 2, 3, 4$ c is in. Letting this procedure continue further yields a binary sequence $(x_1, x_2, ...)$ for each $c \in \mathcal{C}$. Similarly, each of those sequences corresponds to a $c \in \mathcal{C}$. Thus we have a bijection between $\mathcal{C}$ and the binary sequences $(x_i)_{i=1}^{\infty}$. Since the set of binary sequences is uncountable, so is $\mathcal{C}$. $\square$

Remark 2.3.5. An easy way of proving that the Cantor set is uncountable is by using the Hausdorff dimension. It will be shown in the next chapter, that the Hausdorff dimension of every countable set is equal to zero. Moreover, we will show that $\dim_H(\mathcal{C}) = \log_3 2$ and therefore, $\mathcal{C}$ may not be countable.

Lemma 2.3.6. *Let $\mathcal{C}$ be the middle thirds Cantor set. Then*

$$\overline{\dim}_M(\mathcal{C}) = \underline{\dim}_M(\mathcal{C}) = \log_3 2.$$

Proof. Suppose $\mathcal{C}$ is the middle thirds Cantor set. By construction, $\mathcal{C}$ has a covering with 2^n intervals of length $1/3^n$. As a result, for $3^{-n} \leq \varepsilon < 3^{-n+1}$ we have $N(\mathcal{C}, \varepsilon) \leq 2^n$. Hence,

$$\begin{aligned} \overline{\dim}_M(\mathcal{C}) &\leq \limsup_{\varepsilon \to 0} \frac{\log N(\mathcal{C}, \varepsilon)}{\log(1/\varepsilon)} \\ &\leq \frac{\log 2^n}{\log 3^n} = \frac{n \log 2}{n \log 3} = \log_3 2. \end{aligned}$$

Conversely, any interval of length $1/3^n$ can hit $\mathcal{C} \cap I_n$ for at most two of the 2^n nth generation intervals I_n. Hence, $N(\mathcal{C}, \varepsilon) \geq 2^{n-1}$ and thus

$$\begin{aligned} \underline{\dim}_M(\mathcal{C}) &\geq \liminf_{\varepsilon \to 0} \frac{\log N(\mathcal{C}, \varepsilon)}{\log(1/\varepsilon)} \\ &\geq \frac{\log 2^n}{\log 3^n} = \log_3 2. \end{aligned}$$

Therefore, the Minkowski dimension of $\mathcal{C}$ exists and $\dim_M(\mathcal{C}) = \log_3 2$. □

2.3.3 The von Koch curve

We begin with the construction of the von Koch curve which has some similarities to that of the Cantor set. Start with a closed unit interval, as the set of points $\{(x, 0) \in \mathbb{R}^2 : 0 \leq x \leq 1\}$. At the first stage, remove the middle third of the interval and replace it with two line segments of length $1/3$ to make a tent. In other words, cut out the middle third and connect the point $(\frac{1}{3}, 0)$ to $(\frac{1}{2}, \frac{\sqrt{3}}{6})$ with a straight line and do the same with the points $(\frac{1}{2}, \frac{\sqrt{3}}{6})$ and $(\frac{2}{3}, 0)$. The resulting set consists of 4 line segments each of length $1/3$. At the next stage, repeat this procedure on all of the existing line segments, scaled down by a factor of $1/3$.

This results in a set that contains 16 line segments each of length 1/9. The procedure is repeated and at each stage there are 4^n line segments each of length $1/3^n$. When $n \to \infty$, the resulting set is called the Koch curve (see Fig. 2.2).

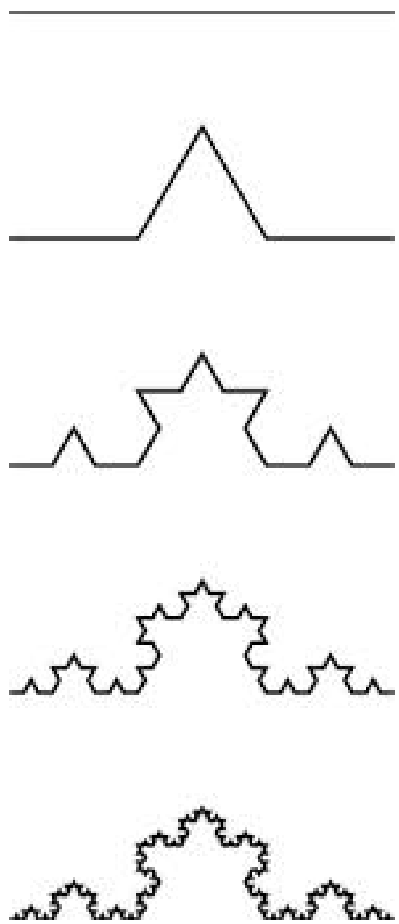


Figure 2.2: The first four iterations of the von Koch curve

Lemma 2.3.7. *Let K be the von Koch curve. Then*

$$\overline{\dim}_M(K) = \underline{\dim}_M(K) = \log_3 4.$$

Proof. Suppose K is the von Koch curve, then by construction, K has a covering with 4^n intervals (and therefore also balls) of length $1/3^n$. Therefore, for $3^{-n} \leq \varepsilon < 3^{-n+1}$ we have

$N(K,\varepsilon) \leq 4^n$. Hence, by Remark 2.1.5 we have

$$\overline{\dim}_M(K) \leq \limsup_{\varepsilon \to 0} \frac{\log N(K,\varepsilon)}{\log(1/\varepsilon)} \leq \frac{\log 4^n}{\log 3^n} = \frac{\log 4}{\log 3}.$$

Moreover, we can see that any interval of length $1/3^n$ intersecting K can intersect at most two of the 4^n nth generation intervals. Hence, we have $N(K,\varepsilon) \geq 4^{n-1}$, which leads to

$$\underline{\dim}_M(K) \geq \liminf_{\varepsilon \to 0} \frac{\log N(K,\varepsilon)}{\log(1/\varepsilon)} \geq \frac{\log 4^n}{\log 3^n} = \frac{\log 4}{\log 3}.$$

Therefore, the Minkowski dimension of K exits and $\dim_M(K) = \log_3 4$. □

2.3.4 Sierpiński carpet

To construct the Sierpiński carpet we begin with the unit square. The square is then divided into 9 identical squares each with sides of length 1/3. Next we remove the central square and subdivide each of the remaining 8 into 9 identical squares, each with sides of length 1/9. Again we remove the central ones. The procedure is then applied recursively to the remaining 8 squares indefinitely (see Fig. 2.3). In other words, let $X = \left\{ \left(\sum_{n=1}^{\infty} \frac{i_n}{3^n}, \sum_{n=1}^{\infty} \frac{j_n}{3^n} \right) : (i_n, j_n) \in S \right\}$ where $S = \{0,1,2\} \times \{0,1,2\} - \{(1,1)\}$.

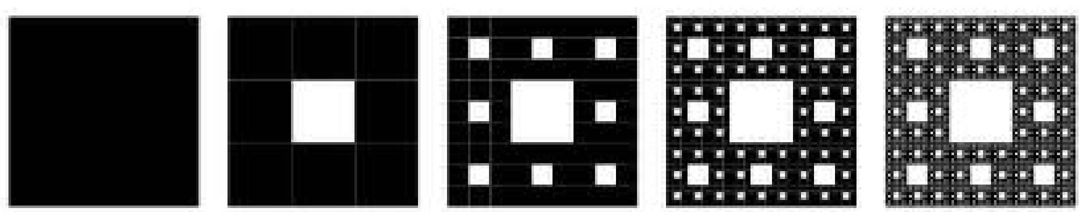


Figure 2.3: The first four iterations of the Sierpiński Carpet

Lemma 2.3.8. *Let X be the Sierpiński carpet. Then*

$$\dim_M(X) = \log_3 8.$$

Proof. Suppose that X is the Sierpiński carpet, then when $\varepsilon_n = \frac{1}{3^n}$ it is possible to cover the set X by 8^n boxes of size $\frac{1}{3^n}$:

$$X_n = \left\{ \left(\sum_{k=1}^{n} \frac{i_k}{3^k} + \frac{s}{3^n}, \sum_{k=1}^{n} \frac{j_k}{3^k} + \frac{t}{3^n} \right) : (i_k, j_k) \in S \text{ and } 0 \leq s, t \leq 1 \right\}.$$

Moreover, it is easy to see that there is no cover with less squares. For any $\varepsilon > 0$ we can choose $\varepsilon_{n+1} \leq \varepsilon < \varepsilon_n$ and we know that $N(X, \varepsilon_n) \leq N(X, \varepsilon) \leq N(X, \varepsilon_{n+1})$. Then

$$\begin{aligned} \frac{n}{(n+1)} \frac{\log 8}{\log 3} = \lim_{\epsilon \to 0} \frac{\log N(X, \varepsilon_{n+1})}{\log(1/\varepsilon_n)} \quad &\leq \quad \lim_{\epsilon \to 0} \frac{\log N(X, \varepsilon)}{\log(1/\varepsilon)} \\ \leq \lim_{\epsilon \to 0} \frac{\log N(X, \varepsilon_n)}{\log(1/\varepsilon_{n+1})} \quad &= \quad \frac{(n+1)}{n} \frac{\log 8}{\log 3}. \end{aligned}$$

Letting $n \to +\infty$, gives us that $\dim_M(X) = \log_3 8$, as desired. $\square$

2.3.5 Two Sierpiński gaskets

We will describe two constructions for the Sierpiński gasket. First is the more familiar one, while the second one will be more suited for certain applications later. The two sets are not the same, however one of the sets can be mapped to the other one by an affine mapping of the plane. This mapping takes the equilateral triangle, with side length equal to 1, to the right triangle with vertices at $0, 1$ and $1 + i$ in the complex plane $\mathbb{C}$.

First construction of the Sirepiński gasket.

To construct the standard Sierpiński gasket we consider a closed unit equilateral triangle, S_0. Start with dividing S_0 into four equally big triangles by joining the midpoint of one side with the other two. Next, remove the middle triangle, i.e. the open triangle containing the center S_0. The remaining set, S_1, consists of three smaller copies of the original triangle, now with the side length of 1/2. Continuing with the same procedure with each of these three triangles leaves us with nine smaller equally big triangles and after that we have 27 smaller triangles and so on. Iterating further in infinitely many steps finally gives us the limit set

$$S = \bigcap_{n=0}^{\infty} S_n,$$

known as the Sierpiński gasket (see Fig.2.4)[7].

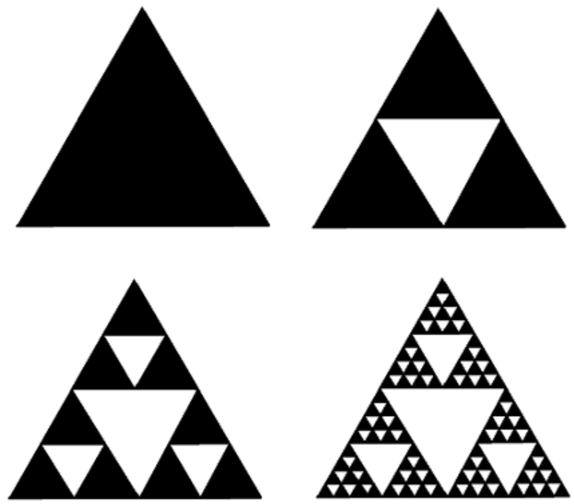


Figure 2.4: The first four iterations of the Sierpiński Gasket

Lemma 2.3.9. *Let S be the Sierpiński gasket. Then*

$$\dim_M(S) = \log_2 3.$$

Proof. To calculate the Minkowski dimension of S we will use Remark 2.1.5 again, this time with triangle shaped coverings. A first covering of S would of course be with a triangle of unit size. The next step is to cover S with three triangles of side $\varepsilon_1 = 1/2$, thus giving us $N(S, \varepsilon_1) = N(S, 1/2) = 3$. Next we cover S with triangles of side $\varepsilon_2 = \frac{\varepsilon_1}{2} = 1/4$ giving us a covering of S with nine triangles of side $1/4$. Following the pattern we have that, in general, S needs to be covered by 3^n triangles of side length 2^{-n}, thus we have

$$N(S, \varepsilon_n) = N(S, (1/2)^n) = 3^n.$$

Thus

$$\begin{aligned}
\overline{\dim}_M(S) = \limsup_{\varepsilon\to 0} \frac{\log N(S,\varepsilon_n)}{\log(1/\varepsilon)} &\leq \lim_{n\to\infty} \frac{\log N(S,\varepsilon_n)}{\log(1/\varepsilon_n)} \\
&\leq \lim_{n\to\infty} \frac{\log N(S,(1/2)^n)}{\log(1/(1/2)^n)} = \lim_{n\to\infty} \frac{\log(3^n)}{\log(2^n)} \\
&= \log_2 3.
\end{aligned}$$

Similarly, calculating the lower Minkowski dimension yields

$$\begin{aligned}
\underline{\dim}_M(S) &\geq \log_2 3 \\
\Longrightarrow \dim_M(S) &= \log_2 3
\end{aligned}$$

which proves our lemma. □

The second construction of the Sierpiński gasket

The following set is constructed in a fashion similar to the Sierpiński carpet. To construct it we again begin with the unit square. The square is then divided into 4 identical squares each with side length 1/2. Next we remove the top left square and subdivide each of the remaining 3 squares into 4 equal sub-squares of side length 1/4. Again we remove the top left ones. The procedure is then applied inductively to the remaining 9 squares indefinitely (see Fig. 2.3). In other words, let

$$\mathcal{SG} = \left\{ \left(\sum_{n=1}^{\infty} \frac{i_n}{2^n}, \sum_{n=1}^{\infty} \frac{j_n}{2^n} \right) : (i_n, j_n) \in S \right\}$$

where now $S = \{0,1\} \times \{0,1\} - \{(0,1)\}$.

Lemma 2.3.10. *Let $\mathcal{SG}$ be the set constructed above. Then*

$$\dim_M \mathcal{SG} = \log_2 3$$

.

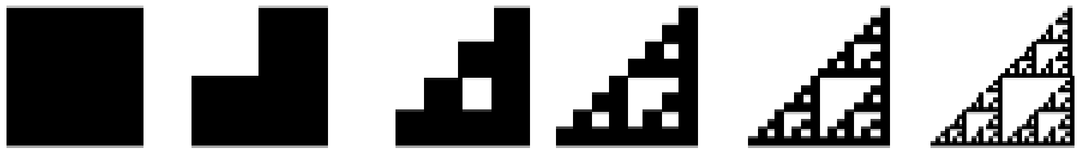


Figure 2.5: The first five iterations of the Sierpiński Gasket

The proof is the same as for the first construction, except for considering coverings with squares instead of triangles.

2.3.6 Self-similar sets

The simplest fractals are self-similar, i.e. they are made of scaled-down copies of themselves, all the way down to arbitrarily small scales[9]. The dimension of such fractals can be defined by extending an elementary observation about classical self-similar sets like line segments, squares, or cubes. In fact, all of the above examples are self-similar sets where the Minkowski dimension and Hausdorff dimension agree. For self-similar sets, there exists a general formula for calculating the dimension, sometimes known as the *similarity dimension*. Suppose a set E consists of m copies of itself, each copy being n-times smaller than the set itself, then the Hausdorff dimension, as well as the Minkowski dimension, are equal to

$$\dim_H(E) = \dim_M(E) = \frac{\log m}{\log n}. \tag{2.3.1}$$

The proof of this result can be found in Falconer's book *Fractal Geometry*[4]. This simple formula is in fact in contrast with what we will see for self-affine sets in Chapter 5. The formula holds under a so called "open set condition" of P.A.P. Moran[11], which is satisfied for all the examples above.

2.4 Sets defined by digit restrictions

In this section, we define a class of sets obtained by restricting the digits which appear in the b-adic expansion of points which belong to the set. One such example is the Middle thirds Cantor set: the collection of points in $[0, 1]$ which do not have any 1's in their ternary expansion. The Sierpiński carpet and gasket are examples of the two dimensional versions of such sets. In addition, the self-affine sets of McMullen and Bedford are also such sets which we will see in Chapter 5.

We will consider a particular example of sets defined by digit restrictions, which will be used later to investigate the vertical slices of the Sierpiński Gasket. We start by recalling some standard definitions.

Definition 2.4.1. The *b-adic expansions* of real numbers for some integer $b \geq 2$ is defined as follows: to each real $x \in [0, 1]$ we have the sequence $\{x_n\} \in \{0, 1, ..., b-1\}^{\mathbb{N}}$ such that

$$x = \sum_{n=0}^{\infty} x_n b^{-n}.$$

Definition 2.4.2. For $S \subset \mathbb{N}$ the *upper density*[1] of S is

$$\overline{d}(S) = \limsup_{N\to\infty} \frac{|S \cap \{1, ..., N\}|}{N}.$$

The *lower density* is

$$\underline{d}(S) = \liminf_{N\to\infty} \frac{|S \cap \{1, ..., N\}|}{N},$$

and if these two values are equal then the limit exist and is simply called $d(S)$, the density of S.

Remark 2.4.3. By restricting the digits we are allowed to use the b-adic expansions which give rise to Cantor sets. In addition, the advantage of the intervals is that they are nested, so either two such intervals are disjoint or one is contained in the other[1].

Example 2.4.4. Suppose $S \subset \mathbb{N}$, and define

$$A_S = \left\{ x \in [0,1], \text{ s.t. } x = \sum_{k=1}^{\infty} \frac{x_k}{2^k} \right\}$$

where

$$x_k \in \begin{cases} \{0,1\} & \text{if } k \in S, \\ \{0\} & \text{if } k \notin S. \end{cases}$$

According to Bishop[1], we can also construct A_S geometrically as follows. Start with the interval $[0,1]$ and subdivide it into two equal length subintervals $[0,1/2]$ and $[1/2,1]$. If $1 \in S$ then keep both intervals and if $1 \notin S$ then keep only the leftmost, $[0,1/2]$. Cut each of the remaining intervals in half, keeping both subintervals if $2 \in S$ and only keeping the left interval otherwise. In general, at the nth step we have a set $A_n^S \subset A_{n-1}^S$ which is a finite union of intervals of length 2^{-n}. We cut each of the intervals in half, keeping both subintervals if $n \in S$ and throwing away the right hand one if $n \notin S$. The limiting set is $A_S = \lim A_n^S$.

Lemma 2.4.5. *Let $S \subset \mathbb{N}$ and A_S be defines as above. Then*

$$\overline{\dim}_M(A_S) = \overline{d}(S), \tag{2.4.1}$$

$$\underline{\dim}_M(A_S) = \underline{d}(S). \tag{2.4.2}$$

Proof of Lemma 2.4.5. The construction of the set S easily implies that S hits exactly as many dyadic intervals of generation n as there are 1's in $(x_1, x_2, \ldots, x_n)$. In other words, that number is equal to

$$2^{\sum_{k=1}^{n} \chi_S(k)}$$

where χ_S is the characteristic function of S, i.e. $\chi_S(n) = 1$ for $n \in S$, and $\chi_S(n) = 0$ for $n \notin S$. So $N(S, 2^{-n})$ is equal to $2^{\sum_{k=1}^{n} \chi_S(k)}$ up to a bounded additive constant. Therefore,

$$\log_2 N(S, 2^{-n}) \approx \sum_{k=1}^{n} \chi_S(k).$$

Thus by the remark made before the proof we obtain

$$\begin{aligned}\overline{\dim}_M(A_S) &= \limsup_{n\to\infty} \frac{\log N(S, 2^{-n})}{\log 2^n} = \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \chi_S(k) \\ &= \limsup_{N\to\infty} \frac{|S \cap \{1, \ldots, N\}|}{N} = \overline{d}(S).\end{aligned}$$

In a similar way, we obtain

$$\underline{\dim}_M(A_S) = \underline{d}(S),$$

which completes the proof □

One interesting feature of this example is that it gives an easy recipe of constructing examples of Cantor sets which have different upper and lower Minkowski dimensions. In particular, sets whose Minkowski dimension do not exist. Indeed, all that is needed is to construct a set S with the property that $\underline{d}(S) < \overline{d}(S)$ and consider the corresponding set A_S. As an illustration, next we construct a Cantor set in the interval $[0, 1]$ that has lower Minkowski dimension equal to 0 while the upper Minkowski dimension is equal to 1. By the previous discussion, we only need the following lemma.

Lemma 2.4.6. *There is a subset of natural numbers $S \subset \mathbb{N}$ such that*

$$\underline{d}_S = 0 \text{ and } \overline{d}_S = 1.$$

Proof. The idea is to consider a subset of integers that has long "gaps" followed by even longer "intervals." For that, let

$$S = \bigcup_{k=1}^{\infty} \{(2k)!, \ldots, (2k+1)!\}.$$

So S consists of intervals of integers of length $(2k+1)! - (2k)!$ followed by a gap of length

$(2k+2)! - (2k+1)!$. Therefore, we have

$$\begin{aligned}
\frac{|S \cap \{1, \ldots, (2k)!\}|}{(2k)!} &= \frac{1}{(2k)!} \sum_{i=1}^{(2k)!} \chi_S(i) = \frac{1}{(2k)!} \left[\sum_{i=1}^{(2k-1)!} \chi_S(i) + \sum_{i=(2k-1)!}^{(2k)!} \chi_S(i) \right] \\
&\leq \frac{1}{(2k)!} \left[\sum_{i=1}^{(2k-1)!} 1 + \sum_{i=(2k-1)!}^{(2k)!} 0 \right] \\
&= \frac{1}{(2k)!} (2k-1)! = \frac{1}{2k} \underset{k \to \infty}{\longrightarrow} 0.
\end{aligned}$$

And therefore $\underline{d}_S = 0$. In a similar way,

$$\begin{aligned}
\frac{|S \cap \{1, \ldots, (2k+1)!\}|}{(2k+1)!} &= \frac{1}{(2k+1)!} \sum_{i=1}^{(2k+1)!} \chi_S(i) = \frac{1}{(2k+1)!} \left[\sum_{i=1}^{(2k)!} \chi_S(i) + \sum_{i=(2k)!}^{(2k+1)!} \chi_S(i) \right] \\
&\geq \frac{1}{(2k+1)!} \left[\sum_{i=1}^{(2k-1)!} 0 + \sum_{i=(2k)!}^{(2k+1)!} 1 \right] \\
&= \frac{1}{(2k+1)!} [0 + (2k+1)! - (2k)!] \\
&= 1 - \frac{(2k)!}{(2k+1)!} = 1 - \frac{1}{2k+1} \underset{k \to \infty}{\longrightarrow} 1.
\end{aligned}$$

Which shows that $\overline{d}_S = 1$. $\square$

The following is an easy consequence of the discussion in this section.

Corollary 2.4.7. *There is a Cantor set $C \subset [0,1]$ such that $\underline{\dim_M}(A_S) = 0$ and $\overline{\dim_M}(A_S) = 1$.*

Proof. By Lemma 2.4.5 and using as S the subset of integers constructed in the last Lemma, we have that

$$\begin{aligned}
\underline{\dim_M}(A_S) &= \underline{d}_S = 0, \\
\overline{\dim_M}(A_S) &= \overline{d}_S = 1.
\end{aligned}$$

In particular we see that Minkowski dimension of A_S does not exist. $\square$

Remark 2.4.8. In the next chapter we consider Hausdorff dimension of sets. In particular it will be shown that $\dim_H(X) \leq \underline{\dim}_M(X)$ for every set $X \subset \mathbb{R}^N$. Therefore, in the last example, even though $\dim_M(A_S)$ doesn't exist, Hausdorff dimension is in fact equal to 0.

Chapter 3

Hausdorff Dimension

As mentioned above, the Minkowski dimension lacks the property of countable stability which gives way to our next notion of dimension, known as Hausdorff. Of all the fractal dimensions known today, Hausdorff dimension is probably the most useful. Based on a construction of Carathéodory, Hausdorff dimension is the oldest known fractal dimension. Most of the primary definitions and results on Hausdorff measures and dimension are due to Constantin Carathéodory and Felix Hausdorff[9]. Hausdorff dimension has the convenience of being defined for any set and is based on measures. However, a big drawback is that it often tends to be quite difficult to calculate or to estimate by computational methods.

3.1 Definitions

Definition 3.1.1. Given any set K in $\mathbb{R}^N$ and $\alpha \geq 0$, we define the *α-dimensional Hausdorff content* of K, by

$$\mathcal{H}^{\alpha}_{\infty}(K) = \inf \left\{ \sum_i |U_i|^{\alpha} : |U_i| < \infty \right\},$$

where $\{U_i\}$ is a countable cover of K by balls, with $|U_i|$ denoting the diameter of U_i as before.

Definition 3.1.2. We say μ is a *measure* on elements in the *σ-algebra*[9] of subsets of $\Omega \subseteq \mathbb{R}^N$ if μ satisfies the following three properties:

1. $\mu(\emptyset) = 0$, i.e. the empty set has zero measure;

2. *Monotonicity*: $\mu(A) \leq \mu(B)$ if $A \subset B \subset \mathbb{R}^N$, and $A, B \in \sigma$-*algebra* ;

3. *Countable sub-additivity*: If $A_1, A_2, ... \subset \mathbb{R}^N$ is a countable (or finite) sequence of sets, then
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i)$$
and
$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$
if A_i are pairwise disjoint Borel sets[4], or $A_i \cap A_j = \emptyset$ for $i \neq j$.

Additionally, we say μ is a *probability measure* if $\mu(\Omega) = 1$.

Definition 3.1.3. Let K be a set in $\mathbb{R}^N$. For every $\alpha \geq 0$ and $\delta > 0$ define

$$\mathcal{H}^{\alpha}_{\delta}(K) = \inf\left\{\sum_{i=1}^{\infty} |U_i|^{\alpha} : |U_i| \leq \delta\right\},$$

where $\{U_i\}$ is a countable cover of K by balls of diameter no more than δ. Then

$$\mathcal{H}^{\alpha}(K) = \sup_{\delta \geq 0} \mathcal{H}^{\alpha}_{\delta}(K) = \lim_{\delta \downarrow 0} \mathcal{H}^{\alpha}_{\delta}(K)$$

is the α-*Hausdorff measure* of the set K, where the limit exits by monotonicity. This gives a measure because it satisfies the properties of a measure defined above in Definition 3.1.2[9].

When $\alpha = 1, 2, 3$, the Hausdorff measure approximates respectively a set's length, area or volume through covers with diameters less than or equal to δ. The approximations become more and more accurate with the smaller sets we use in the coverings. This makes it natural to let $\delta \to 0$ in the definition.

Sometimes the above definitions are given a little differently (see, for instance, Mattila's book *Geometry of Sets and Measure in Euclidean Spaces*[9]). Nonetheless, the various versions differ only by a multiplicative constant. For the calculations of a critical exponent, the

Hausdorff dimension, these differences do not play a role. This is seen from the following lemma.

Lemma 3.1.4. *For* $0 \leq \alpha < \beta < \infty$ *and* $K \subset X$

1. $\mathcal{H}^\alpha(K) < \infty \implies \mathcal{H}^\beta(K) = 0$,

2. $\mathcal{H}^\beta(K) > 0 \implies \mathcal{H}^\alpha(K) = \infty$.

Proof. Fix $\delta > 0$ such that $\mathcal{H}^\alpha_\delta(K) \leq \mathcal{H}^\alpha(K) + 1$. Let $K \subset \bigcup_i E_i$ with $|E_i| \leq \delta$ and $\sum_i |E_i|^\alpha \leq \mathcal{H}^\alpha_\delta(K) + 1$. Then

$$\begin{aligned} \mathcal{H}^\beta_\delta(K) &\leq \sum_i |E_i|^\beta \leq \delta^{\beta-\alpha} \sum_i |E_i|^\alpha && (3.1.1) \\ &\leq \delta^{\beta-\alpha}(\mathcal{H}^\alpha_\delta(K) + 1) && (3.1.2) \end{aligned}$$

which implies that $\mathcal{H}^\beta_\delta(K) \to 0$ since $\beta - \alpha > 0$. We have that $\delta^{\beta-\alpha} \to 0$ as $\delta \downarrow 0$, since $\mathcal{H}^\alpha_\delta(K) < \infty$. Part (2) is a restatement of (1) so the proof would be the same. $\square$

The previous lemma motivates the following definition of Hausdorff dimension.

Definition 3.1.5. The *Hausdorff dimension* of a set $K \subset X \subset \mathbb{R}^N$ is defined to be

$$\begin{aligned} \dim_H(K) &= \sup\{\alpha : \mathcal{H}^\alpha_\infty(K) > 0\}, && (3.1.3) \\ &= \inf\{\beta : \mathcal{H}^\beta_\infty(K) < \infty\}. && (3.1.4) \end{aligned}$$

Remark 3.1.6. The above definition can also be expressed as follows

$$\begin{aligned} \dim_H(K) &= \sup\{\alpha : \mathcal{H}^\alpha(K) = \infty\} \\ &= \inf\{\beta : \mathcal{H}^\beta(K) = 0\} \end{aligned}$$

since $\mathcal{H}^\beta_\infty(K) < \infty \Leftrightarrow \mathcal{H}^\beta(K) < \infty$. In addition, it can also be expressed as

$$\mathcal{H}^\alpha(K) = \begin{cases} \infty & \text{if } \alpha < \dim_H(K) \\ 0 & \text{if } \alpha > \dim_H(K) \end{cases}$$

and $\mathcal{H}^\alpha(K)$ can attain any value in $[0,\infty]$ for $\alpha = \dim_H(K)$.

In other words, $\dim_H(K)$ is the critical value where the α-dimensional Hausdorff measure of the set K essentially jumps from infinity to zero (see Fig. 3.1).

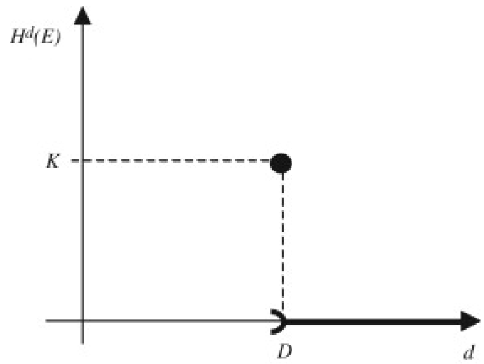


Figure 3.1: Graph of $\mathcal{H}^d(E)$ against d for a set E. The Hausdorff dimension is the value of d at which the 'jump' from ∞ to 0 occurs.

3.2 Basic properties of Hausdorff dimension

Here are some of the main and well known properties of Hausdorff dimension.

Lemma 3.2.1. *Let E, F and F_i be subsets of $\mathbb{R}^N$, then the following properties hold:*

1. ***Monotonicity**: If $E \subset F$ then $\dim_H(E) \leq \dim_H(F)$.*

2. ***Countable Stability**: If $F_1, F_2, \ldots$ is a countable collection of sets, then*

$$\dim_H \bigcup_{i=1}^{\infty} F_i = \sup_{1\leq i<\infty} \{\dim_H(F_i)\}.$$

3. *For all $E \subset \mathbb{R}^N$,*

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

Remark 3.2.2. From the definition of the Hausdorff measure and $N(E,\delta)$ we can deduce the useful relation $\mathcal{H}^\alpha_\delta(E) \leq N(E,\delta)\delta^\alpha$ [7]. In other words, if $\{U_i\}_{i=1}^\infty$ is a countable cover of E, then as $|U_i| \leq \delta$ for all i, we have

$$\sum_{i=1}^{\infty} |U_i|^\alpha \leq \delta^\alpha + \delta^\alpha + \cdots + \delta^\alpha = N(E,\delta)\delta^\alpha.$$

Proof. 1. This is immediate from the measure property that $\mathcal{H}^\alpha(E) \leq \mathcal{H}^\alpha(F)$ for each α and the definition of Hausdorff dimension.

2. It is easy to see that $\dim_H \bigcup_{i=1}^\infty F_i \geq \dim_H(F_j)$ for each j, from the monotonicity property. On the other hand, if $s > \dim_H(F_i) \ \ \forall i$, then $\mathcal{H}^s(F_i) = 0$, such that $\mathcal{H}^s\left(\bigcup_{i=1}^\infty F_i\right) = 0$, giving the opposite inequality.

3. If $\alpha < \dim_H(E)$, then

$$0 < \mathcal{H}^\alpha(E) = \lim_{\delta \downarrow 0} \mathcal{H}^\alpha_\delta(E) \leq \lim_{\delta\downarrow 0} N(E,\delta)\delta^\alpha$$

and thus, $\log N(E,\delta) + \alpha \log \delta > \log \mathcal{H}^\alpha(E) - 1$ for small enough $\delta > 0$. From this, we have

$$0 < \liminf_{\delta\downarrow 0} \frac{\log N(E,\delta) - \log \mathcal{H}^\alpha(E)}{\log(1/\delta)} = \underline{\dim}_M(E) \leq \overline{\dim}_M(E).$$

Then taking the supremum over all α gives us the desired inequality[7].

□

The last property is very useful because it allows us to use the easily computable $\dim_M(E)$ for an upper estimate on $\dim_H(E)$. Generally, Minkowski dimension is easier to calculate because the covering sets are all taken to be of equal size, while Hausdorff incorporates the "weight" of each covering set[7].

3.3 Examples

3.3.1 Countable and open sets

Lemma 3.3.1. *If F is countable, then* $\dim_H(F) = 0$.

Proof. If F_i is a single point then $\dim_H(F_i) = 0$. By countable stability, $\dim_H \bigcup_{i=1} F_i = 0$. $\square$

Lemma 3.3.2. *If $F \subset \mathbb{R}^n$ is open, then $\dim_H(F) = n$.*

Proof. Since F contains a ball of positive n-dimensional volume, $\dim_H(F) \geq n$, but since $F \subset \mathbb{R}^n$, $\dim_H(F) \leq n$ using monotonicity. $\square$

3.3.2 Middle thirds Cantor set: the upper bound

Let us revisit the middle thirds Cantor set $\mathcal{C}$ and calculate its Hausdorff dimension. We start with the estimate from above.

Lemma 3.3.3. *If $\mathcal{C}$ is the middle interval Cantor set then*

$$\dim_H(\mathcal{C}) \leq \log_3 2.$$

Proof. Set $\alpha = \log_3 2$. First we will prove that $\dim_H(\mathcal{C}) \leq \alpha$. We must show that if $\beta > \alpha$ then $\mathcal{H}^\beta(\mathcal{C}) = 0$. Pick $k \geq 0$ and let $I_{0,1}, ..., I_{k,2^k}$ be the 2^k intervals that comprise $\mathcal{C}_k$, each of length $1/3^k$ in the construction of the Cantor set from Section 2.3.2 . Since $\mathcal{C} \subseteq \mathcal{C}_k$, this is a cover of $\mathcal{C}$. We compute the β-length of the cover. It follows that

$$\begin{aligned} \sum_{j=1}^{2^k} |I_{k,j}|^\beta &= \sum_{j=1}^{2^k} (3^{-k})^\beta = (2^k)(3^{-\beta k}) \\ &= \left(\frac{2}{3^\beta}\right)^k. \end{aligned}$$

Since $\beta > \log_3 2$, we have $2/3^\beta < 1$ and

$$\left(\frac{2}{3^\beta}\right)^k \xrightarrow[k \to \infty]{} 0.$$

Therefore, $\mathcal{H}^\beta(\mathcal{C}) = 0$ for every $\beta > \alpha$ we obtain that $\dim_H(\mathcal{C}) \leq \alpha = \log_3 2$. $\square$

Next, we want to determine the lower bound to show that $\dim_H(\mathcal{C}) \geq \alpha = \log_3 2$. To do this, we will need to introduce a new technique known as the Mass Distribution Principle.

3.3.3 Von Koch snowflake, Sierpiński carpet and Sierpiński gasket

In a way which is similar to the proof of the upper bound for the Hausdorff dimension of the Cantor set, we can also obtain upper bounds for the Hausdorff dimension of the Koch Snowflake, Sierpiński carpet and gasket.

Lemma 3.3.4. *Let S, C and G be the Koch Snowflake, Sierpiński carpet and Sierpiński gasket, respectively. Then*

$$\begin{aligned}
\dim_H(K) &\leq \log_3 4, \\
\dim_H(C) &\leq \log_3 8, \\
\dim_H(G) &\leq \log_2 3.
\end{aligned}$$

The idea of the proof in all the cases (like in the case of the Cantor set) is to construct the appropriate covering of the corresponding sets. In fact in all these cases one may take the "canonical" coverings which are given in the construction. We omit the details of the proofs of these inequalities and refer the interested reader to Falconer's book[5].

An indirect proof of all these inequalities comes from part 3 of Lemma 3.2.1: $\dim_H(E) \leq \dim_M(E)$ for every set $X \subset \mathbb{R}^N$.

3.4 Lower estimates of Hausdorff dimension

Since every covering of a set E gives an upper bound for $\dim_H(E)$, upper bounds are easier to compute than lower bounds. Lower bounds for Hausdorff dimension are conventionally obtained by constructing an appropriate measure supported on the set from the so called Mass Distribution Principle. In the next chapter, we will generalize the Mass Distribution Principle by proving Billingsley's lemma (Theorem 4.1.3).

3.4.1 Mass Distribution Principle

A measure μ on a bounded subset A of $\mathbb{R}^N$ for which $0 < \mu(A) < \infty$ will be called a *mass distribution*[4] on A. One may think of $\mu(A)$ as the mass of the set A. Informally speaking, we take a finite and positive mass and spread it in some way across a set A to get a mass distribution on A.

Theorem 3.4.1. (Mass Distribution Principle): *Suppose E is a subset of $\mathbb{R}^N$ and $\alpha \geq 0$. If there is a non-trivial mass distribution μ on E, i.e. $\mu(E) > 0$, and a constant $0 < C < \infty$ such that*

$$\mu(B(x, r)) \leq Cr^\alpha,$$

for all balls $B(x, r)$ with $x \in \mathbb{R}^N$ and $r > 0$, then

$$\mathcal{H}^\alpha(E) \geq \mathcal{H}^\alpha_\infty(E) \geq \frac{\mu(E)}{C} > 0,$$

and hence, $\dim_H(E) \geq \alpha$.

Proof. Suppose that $U_1, U_2, ...$ is a cover of E by balls with $|U_i| \leq \delta$. For $r_1, r_2, ...$ where $r_i > |U_i|$, consider the cover where we choose x_i in each U_i and take open balls $B(x_i, r_i)$. Then by assumption,

$$\mu(U_i) \leq Cr_i^\alpha.$$

We deduce that $\mu(U_i) \leq C|U_i|^\alpha$, that is,

$$\sum_i |U_i|^\alpha \geq \sum_i \frac{\mu(U_i)}{C} \geq \frac{\mu(E)}{C},$$

which is true from the properties of sub-additivity and monotonicity. Thus,

$$\mathcal{H}^\alpha(E) \geq \mathcal{H}^\alpha_\infty(E) \geq \frac{\mu(E)}{C} > 0.$$

Therefore, $\dim_H(E) \geq \alpha$, as desired[1]. □

3.4.2 Middle thirds Cantor set: lower bound

Lemma 3.4.2. *Let $\mathcal{C}$ be the middle thirds Cantor set. Then*

$$\dim_H(\mathcal{C}) \geq \log_3 2.$$

Proof. We will show that there exists a measure μ on $\mathcal{C}$ such that there is a constant $C < \infty$ such that for every $I \subset \mathbb{R}$ we have $\mu(I) \leq C|I|^{\log_3 2}$.

We will proceed in two steps. First, we will show that $\mu(I_{k,j}) \leq C|I_{k,j}|^{\log_3 2}$ for every k and j. Then we will generalize this and show that the same inequality holds true for any I, possibly with a different constant C.

To define μ we let $\mu([0,1]) = \mu(I_{0,1}) = 1$. Next, we let

$$\mu(I_{1,1}) = \mu(I_{1,2}) = \frac{\mu(I_{0,1})}{2} = \frac{1}{2}.$$

To define μ in general we proceed by induction and let $\mu(I_{k,j}) = \frac{1}{2^k}$, for every $k > 1$ and every $j \in \{1, \dots, 2^k\}$.

Recall from our construction in Section 2.3.2 that we have $|I_{k,j}| = \frac{1}{3^k}$. Therefore, we have

$$\begin{aligned} \frac{1}{2^k} &\leq C\left(\frac{1}{3^k}\right)^{\log_3 2} = C\left(\frac{1}{3}\right)^{k\log_3 2} = C\left[\left(\frac{1}{3}\right)^{\log_3 2}\right]^k \\ &= C\left(\frac{1}{2}\right)^k. \end{aligned}$$

We can take $C = 1$ and we obtain $\mu(I_{k,j}) \leq |I_{k,j}|^{\log_3 2}$, as desired.

To show that in general, $\mu(I) \leq C|I|^{\log_3 2}$, we will choose k so that $\frac{1}{3^k} \leq |I| < \frac{1}{3^{k-1}}$. Since I intersects at most one $I_{k,j}$, this implies that

$$\mu(I) \leq \mu(I_{k,j}) \leq |I_{k,j}|^{\log_3 2}.$$

Therefore, I can intersect at most two intervals of the form $I_{k-1,j}$. If $|I| < \frac{1}{3^{k-1}}$, then I

intersects at most one interval of the from $I_{k-1,j}$. Hence, we have $\mu(I) \leq 2\mu(I_{k-1,j})$ and using the estimate just obtained, we get

$$\begin{aligned} \mu(I) &\leq 2|I_{k-1,j}|^{\log_3 2} \leq 2\left(\frac{1}{3^{k-1}}\right)^{\log_3 2} = 2\left(\frac{3}{3^k}\right)^{\log_3 2} \\ &= 2\frac{3^{\log_3 2}}{(3^k)^{\log_3 2}} = 4 \cdot \left(\frac{1}{3^k}\right)^{\log_3 2} \leq 4 \cdot |I|^{\log_3 2}. \end{aligned}$$

Therefore, we have $\mu(I) \leq 4|I|^{\log_3 2}$ for every $I \subset \mathbb{R}$ and $\dim_H(\mathcal{C}) \geq \log_3 2$. □

Remark 3.4.3. Combining this with the upper bound obtained in 3.3.3, we conclude that

$$\dim_H(\mathcal{C}) = \log_3 2.$$

Chapter 4

Dimension of Measures and Its Applications

4.1 Motivation: Billingsley's Lemma

Now we introduce a more elegant adaptation of the Mass Distribution Principle known as Billingsley's lemma. The difference here is that the measure in Billingsley's lemma needs to satisfy estimates in a neighborhood of each point, where the size of that neighborhood can vary from point to point. In the Mass Distribution Principle, the measure had to satisfy uniform estimates on all balls.

Definition 4.1.1. First we define the class of *Borel sets*[4] as the smallest collection of subsets of $\mathbb{R}^N$ with the following properties:

1. every open set and every closed set is a Borel set;

2. the union of every finite or countable collection of Borel sets is a Borel set, and the intersection of every finite or countable collection of Borel sets is a Borel set.

Remark 4.1.2. Almost all of the subsets of $\mathbb{R}^N$ that will be of any interest will be Borel sets. Any set that can be constructed using a sequence of countable unions or intersections starting with the open sets or closed sets will be Borel. See Falconer's *Fractal Geometry*[4] for the definitions of open and closed sets for the topology on $\mathbb{R}^N$. In fact, we will only look at G_δ or F_σ types of sets.

For Billingsley's lemma, first we assume we have an integer $b \geq 2$ and for $x \in [0,1]$ let $I_n(x)$ denote the n-th generation b-adic interval containing x.

Theorem 4.1.3. (Billingsley's Lemma): *Let $A \subset [0,1]$ be a Borel set and let μ be a finite Borel measure on $[0,1]$. Suppose $\mu(A) > 0$. If*

$$\alpha_1 \leq \liminf_{n\to\infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|} \leq \beta_1,$$

for all $x \in A$, then $\alpha_1 \leq \dim_H(A) \leq \beta_1$.

Proof. Let $\alpha < \alpha_1 < \beta_1 < \beta$. Then the assumptions imply that

$$\limsup_{n\to\infty} \frac{\mu(I_n(x))}{|I_n(x)|^\beta} \geq 1, \tag{4.1.1}$$

and

$$\limsup_{n\to\infty} \frac{\mu(I_n(x))}{|I_n(x)|^\alpha} \leq 1. \tag{4.1.2}$$

We will show that

$$(4.1.1) \Longrightarrow \mathcal{H}^\beta(A) \leq \mu(A) \tag{4.1.3}$$

and

$$(4.1.2) \Longrightarrow \mathcal{H}^\alpha(A) \geq \mu(A). \tag{4.1.4}$$

This would imply the assertion and prove our lemma.

First we prove (4.1.3). For each $0 < c < 1$ and $\epsilon > 0$ we have that for every $x \in A$ there exist an $n(x) \in \mathbb{N}$ such that

$$\frac{\mu(I_{n(x)}(x))}{|I_{n(x)}(x)|^\beta} > c \text{ and } b^{-n(x)} < \epsilon$$

We have that $\{I_{n(x)}(x)\}_{x \in A}$ is a cover of A that has a subcover $\{C_k\}$ of disjoint intervals. This covering of A has the property that $|C_k| < \epsilon$ for all k and

$$\begin{aligned}\sum_k |C_k|^\beta &\leq c^{-1} \sum_k \mu(C_K) \\ &\leq c^{-1}\mu(A).\end{aligned}$$

This gives us

$$\mathcal{H}^\beta_\epsilon(A) \leq c^{-2}\mu(A).$$

Then taking $c \to 0$ and $\epsilon \to 0$ we get (4.1.3). Thus, (4.1.1) implies (4.1.3), as desired.

Next we prove (4.1.4), the second assertion. For $C > 1$ and $m \in \mathbb{N}$, set

$$A_m := \{x \in A : \mu(I_n(x)) < C|I_n(x)|^\alpha, \ \forall\, n > m\}.$$

Since we have $A_m \subset A_{m+1}$ and $A = \bigcup_m A_m$, we have $\mu(A) = \lim_{m\to\infty} \mu(A_m)$. Thus it suffices to prove the estimate for A_m.

Now fix $m \in \mathbb{N}$ such that $\epsilon < b^{-m}$ and consider any cover of A_m by b-adic intervals $\{C_k\}$ with $|C_k| < \epsilon$. Then

$$\begin{aligned}\sum_k |C_k|^\alpha \geq \sum_{k: C_k \cap A_m \neq \emptyset} |C_k|^\alpha &\geq C^{-1} \sum_{k: C_k \cap A_m \neq \emptyset} \mu(C_k) \\ &\geq C^{-1}\mu(A_m).\end{aligned}$$

This gives us

$$\mathcal{H}^\alpha_\epsilon(A) \geq c^{-1}\mu(A_m).$$

Taking $c \to 0$, $\epsilon \to 0$ and $m \to \infty$ we get (4.1.4). Therefore, (4.1.2) implies (4.1.4) which completes our proof. $\square$

4.2 Dimension of a measure

The discussion in the previous sections motivates (at least) two definitions for the dimension of a measure on a set.

4.2.1 Dimension of μ through Billingsley's Lemma

Billingsley's lemma motivates the local study of sets and measures on a set E and may be considered to be a motivation for the study of the dimension of measures[13]. It says that the dimension of a set on which a measure μ is supported depends on the asymptotics of the quantity $\frac{\log \mu(I_n(x))}{\log |I_n(x)|}$ as $n \to \infty$, i.e. on infinitesimal scales around $x \in E$. Recall, that here $I_n(x)$ denotes the b-adic interval of the form $[\frac{j-1}{b^n}, \frac{j}{b^n})$ containing x, where b is any positive integer.

In fact, let us define the *local dimension of* μ *around* x as follows

$$\alpha_\mu(x) = \liminf_{n\to\infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|}.$$

Now, recall that for any two sets $E, F \subset \mathbb{R}^N$ we have

$$\dim_H(E \cup F) = \max\{\dim_H(E), \dim_H(F)\},$$

which means that Hausdorff dimension may be thought of as the dimension of the "largest part" of the set. Therefore, we may introduce the following notion, which may be thought of as "dimension of μ".

Definition 4.2.1. For a set $E \subset \mathbb{R}^N$ and a measure μ on E we define

$$\alpha_\mu = \operatorname{ess\,sup}_E \alpha_\mu(x) = \operatorname{ess\,sup}_E \left\{ \liminf_{n\to\infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|} \right\}. \tag{4.2.1}$$

Remark 4.2.2. Recall that if f is a real-valued and measurable function on a set E of positive measure then the *essential supremum*[19] of f on E is defined as follows: If $|\{x \in E : f(x) > \alpha\}| > 0$ for all $\alpha \in \mathbb{R}$, let $\operatorname{ess\,sup}_E f = +\infty$; otherwise, let

$$\operatorname{ess\,sup}_E f = \inf\{\alpha : |\{x \in E : f(x) > \alpha\}| = 0\}.$$

4.2.2 Dimension of μ through dimension of support

Informally speaking, the dimension of a measure μ is the dimension of a set on which it is supported. Note however, that μ may vanish on large parts of E and thus will not contain any information about those parts. Therefore, it is natural to consider the "smallest" set on which μ is "nontrivial" in some sense. This motivates the following definition.

Definition 4.2.3. (Dimension of a measure): If μ is a measure on $\mathbb{R}^N$ we define

$$\dim(\mu) = \inf\{\dim(A) : \mu(A^c) = 0, \emptyset \neq A \subset \mathbb{R}^N\}. \tag{4.2.2}$$

Remark 4.2.4. An equivalent definition is given by

$$\dim(\mu) = \inf\{\alpha : \mu \perp \mathcal{H}^\alpha\}$$

where $\mu \perp \nu$ means the two measures are mutually singular, i.e. there exists a set $A \subset \mathbb{R}^N$ such that $\mu(A) = \nu(A^c) = 0$.

4.2.3 Dimension of μ as the supremum of local dimension

Interestingly enough the two approaches to the dimension give the same answer.

Theorem 4.2.5. *For a set $E \subset \mathbb{R}^N$ and a measure μ on E we have*

$$\dim(\mu) = \alpha_\mu.$$

For details of the proof and further discussion of "local dimension", see Bishop and Peres' book[1] or Pesin's monograph[13].

4.2.4 Examples

Example 4.2.6. Suppose that the measure μ on the middle thirds Cantor set $\mathcal{C}$ gives equal mass to each nth generation interval in the construction. If we consider a 3-adic interval I of length $1/3^n$ then

$$\mu(I) = 2^{-n} = |I|^{\log_3 2},$$

if I hits $\mathcal{C}$ and is 0 otherwise. Thus

$$\lim_{n\to\infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|} = \log_3 2,$$

for all $x \in \mathcal{C}$ and hence for μ almost every x. Then by the previous theorem, $\dim(\mu) = \log_3 2$, as expected.

We will have two more examples, where calculation of dimension of measures will be used to find Hausdorff dimension of sets. The first example is the dimension of certain product (Bernoulli) measures, which will be used to show the well known fact: "in a random sequence of 0's and 1's either of the digits appears half of the time". This fact in turn will be used to find the dimension of random slices of Sierpiński gasket. The second application of dimension of measures is the computation of Hausdorff dimension of certain self-affine sets, to which the last chapter of this work is dedicated. For both of these examples we will need the Law of Large Numbers.

4.3 Frequency of 1's in a random binary sequence

In this section we apply the techniques developed above to calculate Hausdorff dimension (and measure) of a well known example. The results of this section may be thought of as generalization of the following fact: *In a random binary sequence* 1*'s appear half of the time.*

To give a precise formulation of this statement we need some notation.

Let $p \in (0,1)$ be any real number and define the $A_p \subset [0,1]$ as follows

$$A_p := \left\{ x = \sum_{n=1}^{\infty} \frac{x_n}{2^n} : \lim_{j\to 0} \frac{1}{j} \sum_{k=1}^{j} x_k = p \right\}. \tag{4.3.1}$$

Thus A_p is the set of real numbers $x \in [0,1]$ in which 1 occurs in the binary expansion of x about a pth of the time. For regular real numbers we expect a 1 to occur about half the time. This is indeed the case, the set $A_{1/2}$ is a set of full measure in $[0,1]$.

Corollary 4.3.1. *If $\mathcal{L}^1$ is the Lebesgue measure on $\mathbb{R}$ and $A_{1/2}$ is defined as above then*

$$\mathcal{L}^1(A_{1/2}) = 1. \tag{4.3.2}$$

This is a particular case of the theorem which we formulate next.

Now for $p \in (0,1)$ we define a probability measure μ_p on $[0,1]$ as follows, see [1]. Start by assigning mass 1 to the unit interval $[0,1]$. In the next step, assign mass p to $[0,1/2]$ and mass $(1-p)$ to $[1/2,1]$. To continue, assume that μ_p has been defined on all the dyadic intervals up to generation n. To define the measure on generation n dyadic intervals, pick such an interval J and assume $I = I_{n-1,j}$ is its "parent", i.e. the unique interval of generation $n-1$ containing J. Then let $\mu_p(J) = p\mu_p(I)$ if J is the "left" subinterval of I and let $\mu_p(J) = (1-p)\mu_p(I)$ if J is the "right" subinterval of I.

Since "left" and "right" intervals correspond to a 0 or 1 respectively in the binary expansion of real numbers, we can alternatively define μ_p simply be the following formula

$$\mu_p(I_n(x)) = p^{\sum_{k=1}^{n} x_k}(1-p)^{n-\sum_{k=1}^{n} x_k}, \tag{4.3.3}$$

where $x = \sum_{k=1}^{\infty} x_k 2^{-k}$. Note that $\mu_{1/2}$ is the Lebesgue measure on $[0,1]$.

Lemma 4.3.1 above is the particular case of the following result.

Theorem 4.3.2. *With the definitions above we have that for every $p \in (0,1)$*

$$\mu_p(A_p) = 1. \tag{4.3.4}$$

Proof. To see this, let $f_n(x) = x_n - p$, i.e.

$$f_n(x) = \begin{cases} 1-p & \text{if } x_n = 1, \\ -p & \text{if } x_n = 0. \end{cases}$$

Now, if $S_n = \sum_{k=1}^n f_n$, then unwinding definitions shows that A_p is exactly the set where $\frac{1}{n}S_n \to 0$.

Indeed,

$$\begin{aligned} \frac{1}{j}\sum_{i=1}^{j} x_i \to p \iff \frac{1}{j}\sum_{i=1}^{j} x_i - p\frac{j}{j} \to 0 \quad &\iff \quad \frac{1}{j}\sum_{i=1}^{j} x_i - \frac{1}{j}\sum_{i=1}^{j} p \to 0 \\ &\iff \quad \frac{1}{j}\sum_{i=1}^{j} (x_i - p) \to 0 \\ &\iff \quad \frac{1}{j}\sum_{i=1}^{j} f_i(x) \to 0. \end{aligned}$$

Therefore to prove the theorem, we need to show that

$$\mu_p \left\{ x \in [0,1] : \frac{1}{n}S_n(x) \to 0 \right\} = 1.$$

We will achieve this by applying the Law of Large Numbers, which can be found in Appendix A. To do this we must check that the conditions of the Law of Large Numbers are satisfied. First, we must show $\int |f_i|^2 d\mu_{\frac{1}{2}} \leq 1$. Recall that $|f_i| \leq 1$, then

$$\int_0^1 |f_n|^2 d\mu_p \leq \int_0^1 1^2 d\mu_p = 1\mu_p([0,1]) = 1.$$

Next, we check orthogonality, i.e.

$$\int_0^1 f_i(x) f_j(x) d\mu_p = 0 \text{ if } i \neq j.$$

To check this let $m, n \in \{0, 1\}$ and denote

$$E_{i,j}^{m,n} = \{x \in [0,1] : x_i = m, x_j = n\}.$$

Then we obtain

$$\begin{aligned} \int_0^1 f_i(x) f_j(x) d\mu_p &= \int_0^1 (x_i - p)(x_j - p) d\mu_p \\ &= (1-p)^2(\mu_p(E_{i,j}^{1,1})) - p(1-p)(\mu_p(E_{i,j}^{1,0}) + \mu_p(E_{i,j}^{0,1})) + p^2(\mu_p(E_{i,j}^{0,0})), \end{aligned}$$

Note, that since μ_p is a probability product measure, we will have

$$\mu_p(E_{i,j}^{m,n}) = \mu_p(E_i^m)\mu_p(E_j^n),$$

where $E_i^m = \{x \in [0,1] : x_i = m\}$. From the definitions, we obtain that for every $i \in \mathbb{N}$ we have $\mu(E_i^1) = p$, and $\mu(E_i^0) = 1 - p$, and therefore

$$\begin{aligned} \mu_p(E_{ij}^{11}) &= p^2, \\ \mu_p(E_{ij}^{10}) &= \mu_p(E_{ij}^{01}) = p(1-p), \\ \mu_p(E_{ij}^{00}) &= (1-p)^2. \end{aligned}$$

Thus,

$$\int_0^1 f_i(x) f_j(x) d\mu_p = (1-p)^2 \cdot p^2 - p(1-p) \cdot (2p(1-p)) + p^2(1-p)^2 = 0,$$

which proves the orthogonality of f_i's and therefore the theorem. $\square$

Since $A_{1/2}$ has full Lebesgue measure in the interval it follows that $\dim_H(A_{1/2}) = 1$, while A_p has zero measure for every $p \neq 1/2$. Next we use the previous theorem to calculate the dimension of A_p for every $p \in [0,1]$.

Theorem 4.3.3. *For every $p \in [0,1]$ we have*

$$\begin{aligned}\dim(A_p) = \dim(\mu_p) &= h_2(p) \\ &= -p\log_2 p - (1-p)\log_2(1-p).\end{aligned}$$

The quantity h_p is called the *entropy*[1] of p and is strictly less than 1, except for $p = 1/2$. It represents the uncertainty associated to the probability: if $p = 0$ or 1 the entropy is 0; it is maximized when $p = 1/2$.

Proof. Since $\mu_p(A_p) = 1$, from the definition of a dimension of measures, we have that

$$\dim_H(A_p) \geq \dim(\mu_p).$$

On the other hand, by Theorem 4.2.5, we have that

$$\dim(\mu_p) = \operatorname{ess\,sup}\left\{\liminf_{n\to\infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|}\right\}.$$

Now, by equation (4.3.3) we have

$$\begin{aligned}\frac{\log \mu(I_n(x))}{\log |I_n(x)|} &= \frac{\log(p^{\sum_{k=1}^n x_k}(1-p)^{n-\sum_{k=1}^n x_k})}{\log 2^{-n}} \\ &= -\frac{1}{\log 2}\left\{(\frac{1}{n}\sum_{k=1}^n x_k)\log p + (\frac{1}{n}(n - \sum_{k=1}^n x_k))\log(1-p)\right\}.\end{aligned}$$

Since μ_p a.e. point is in A_p, we have that for μ_p a.e. x gives us that

$$\frac{1}{n}\sum_{k=1}^n x_k \underset{n\to\infty}{\longrightarrow} p$$

and therefore

$$\dim_H(A_p) \geq \dim(\mu_p) = -p\log_2 p - (1-p)\log_2(1-p).$$

To obtain the opposite inequality note that

$$A_p \subset \left\{ x \in [0,1] : \liminf_{n \to \infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|} \right\} = h_2(p),$$

and $\mu_p(A_p) = 1 > 0$. Hence by Billingsley's lemma we get $\dim_H(A_p) \le h_2(p)$. $\square$

4.4 Marstrand's theorem

Suppose that $A \subset \mathbb{R}^2$ has dimension $\dim_H(A)$. Let $L_x = \{(x, y) : y \in \mathbb{R}\}$ be a vertical line. We can make the following assertion about the dimension of a typical intersection $A \cap L_x$.

The next theorem shows that *if the set is large enough then typical slices have dimensions that drop by at least* 1[14].

Theorem 4.4.1. (Marstrand Slicing Theorem): *Let* $A \subset \mathbb{R}^2$ *and suppose* $\dim_H(A) \ge 1$. *Then*

$$\dim_H(A \cap L_x) \le \dim_H(A) - 1,$$

for almost every $x \in \mathbb{R}$.

First we consider and prove the following lemma:

Lemma 4.4.2. *For* $1 \le \alpha \le 2$ *we can write*

$$\mathcal{H}^{\alpha}(A) \ge \int \mathcal{H}^{\alpha-1}(A \cap L_x) dx$$

It suffices to prove this claim: for $\alpha > \dim_H(A)$,

$$0 = \mathcal{H}^{\alpha}(A) \ge \int_{\mathbb{R}} \mathcal{H}^{\alpha-1}(A \cap L_x) dx$$

implies $\mathcal{H}^{\alpha-1}(A \cap L_x) = 0$ for almost every x, i.e. $\dim_H(A \cap L_x) \le \alpha - 1$ for almost every $x \in \mathbb{R}$.

Proof. Given $\varepsilon > 0$ and $\delta > 0$, let $\{U_i\}$ be an open cover of A with $|U_i| < \varepsilon$ such that

$$\sum_i |U_i|^\alpha < \mathcal{H}^\alpha_\varepsilon(A) + \delta.$$

We can cover each U_i by a square Q_i with sides parallel to the axes and with side length $s_i = |U_i|$. Let I_i be the projection of Q_i onto the x-axis. For each x, the slices $\{(Q_i)_x\}$ form a cover of $A \cap L_x$ and have length

$$|(Q_i)_x| = \begin{cases} s_i & \text{if } x \in I_i, \\ 0 & \text{if } x \notin I_i. \end{cases}$$

We now have an ε-cover of $A \cap L_x$ and

$$\mathcal{H}^{\alpha-1}_\varepsilon(A \cap L_x) \leq \sum_i |(Q_i)_x|^{\alpha-1} = \sum_{i:x\in I_i} s_i^{\alpha-1}.$$

Therefore,

$$\begin{aligned} \int_{\mathbb{R}} \mathcal{H}^{\alpha-1}_\varepsilon(A \cap L_x)dx &\leq \int_{\mathbb{R}} \left(\sum_{i:x\in I_i} s_i^{\alpha-1} \right) dx = \sum_i s_i^\alpha \\ &\leq \mathcal{H}^\alpha_\varepsilon(A \cap L_x) + \delta. \end{aligned}$$

Letting $\delta \to 0$ gives

$$\int_{\mathbb{R}} \mathcal{H}^{\alpha-1}_\varepsilon(A \cap L_x)dx \leq \mathcal{H}^\alpha_\varepsilon(A \cap L_x).$$

As $\varepsilon \to 0$, $\mathcal{H}^{\alpha-1}_\varepsilon(A \cap L_x) \nearrow \mathcal{H}^{\alpha-1}(A \cap L_x)$, so the Monotone Convergence Theorem[15] implies

$$\int_{\mathbb{R}} \mathcal{H}^{\alpha-1}(A \cap L_x)dx \leq \mathcal{H}^\alpha(A).$$

This completes the proof of the lemma. □

Proof. (Theorem 4.5.1): Let $\alpha > \dim_H(A)$ then by Lemma 4.4.2

$$0 = \mathcal{H}^\alpha(A) = \int_{-\infty}^{\infty} \mathcal{H}^{\alpha-1}(A \cap L_x)dx.$$

Thus, by Fubini's Theorem[19], $\mathcal{H}^{\alpha-1}(A \cap L_x) = 0$ for almost every $x \in \mathbb{R}$. In particular, $\dim_H(A \cap L_x) \le \alpha - 1$ for such x, as required[12]. □

4.5 Vertical slices of self-similar sets

A natural question arises from Marstrand's theorem. Is it possible to replace the inequality in the theorem by an equality?

For that, we will look at the example of the Sierpiński Gasket and will calculate the Hausdorff dimension of almost all slices by vertical lines. The main technical tool in the next Lemma is the Law of Large Numbers whose proof can be found in Appendix A.

Lemma 4.5.1. *Let G be the Sierpiński Gasket (see Fig. 4.1) and let L_x be the vertical line $L_x = \{(x, y) : y \in \mathbb{R}\}$. For almost every $x \in [0, 1]$, the Minkowski dimension of $L_x \cap G$ is equal to $1/2$,*

$$\dim_M(L_x \cap G) = \frac{1}{2}, \ \ a.e. \ x \in [0, 1].$$

Remark 4.5.2. Since all the vertical slices of the Sierpiński Gasket are very well behaved Cantor sets, we could use Billingsley's lemma to show that for these slices the Hausdorff and Minkowski dimensions are equal to each other. In fact, these sets are examples of "sets defined by digit restrictions" defined in Chapter 2.

Remark 4.5.3. Recall that $\dim_M(G) = \dim_H(G) = \frac{\log 3}{\log 2} \approx 1.58496250072$. Therefore

$$\frac{1}{2} < \dim_H(G) - 1 \approx .58496...$$

and thus it is impossible to replace the inequality by an equality in Marstrand's theorem.

Proof. Let $x \in [0, 1]$ and let $(x_1, x_2, \dots)$ be its expansion. Let $\sigma_k(x)$ be defined as the number of 1's in the binary expansion of x up to k, also let $C_x := L_x \cap G$.

Before calculating the Hausdorff dimension of C_x let us study how it may be constructed from the binary expansion of x.

Construction of C_x is by induction. If $x_1 = 0$ then C_x is contained in the lower half of $\{x\} \times [0,1]$. In other words, if $x_1 = 0$ then we divide $\{x\} \times [0,1]$ into two equal intervals of length $1/2$ and remove the top one. On the other hand if $x_1 = 1$ then we divide $\{x\} \times [0,1]$ into two equal intervals and keep both. Similarly, in the second step if $x_2 = 0$ we divide all the intervals of length $1/2$ obtained in the first step to two equal length intervals of length $1/4$ and remove the top halves. If $x_2 = 1$, we divide all the intervals left from the first step to two equal parts and keep both. By induction, assume that the k-th approximation to C_x, which we denote by C_x^k has been constructed, i.e. we have chosen all the k-th generation dyadic subintervals in $\{x\} \times [0,1]$ of length 2^{-k}, whose interiors have a non-empty intersection with C_x. To construct C_x^{k+1}, i.e. the $(k+1)$-th approximation of C_x we divide all the intervals remaining from the previous step to two equal subintervals and remove the top ones if $x_{k+1} = 0$ or keep both if $x_{k+1} = 1$. Finally we let

$$C_x = \bigcap_{k=1}^{\infty} C_x^k.$$

Next, we need to estimate $N(C_x, 2^{-k})$. For that, set

$$n_i(x) = \begin{cases} 1 & \text{if } x_i = 0, \\ 2 & \text{if } x_i = 1. \end{cases}$$

Then from the construction described above, we clearly have that C_x may be covered by $\prod_{i=1}^{k} n_i(x)$ dyadic intervals of length 2^{-k}. Just like for the sets defined by digit restrictions, we obtain

$$\begin{aligned} N(C_x, 2^{-k}) &\approx n_1(x) \cdot n_2(x) \cdot ... \cdot n_k(x) \\ &= 1^{k-\sigma_k(x)} \cdot 2^{\sigma_k(x)} \\ &= 2^{\sigma_k(x)}, \end{aligned}$$

where " $\approx$ " means "up to an additive constant".

Thus, for every $x \in [0,1]$ we have

$$\begin{aligned}
\overline{\dim}_M(C_x) &= \limsup_{k\to\infty} \frac{\log N(C_x, 2^{-k})}{\log 2^k} \\
&= \limsup_{k\to\infty} \frac{\log 2^{\sigma_k(x)}}{k \log 2} = \limsup_{k\to\infty} \frac{\sigma_k(x)}{k} \log_2 2 \\
&= \limsup_{k\to\infty} \frac{\sigma_k(x)}{k}.
\end{aligned}$$

and similarly

$$\underline{\dim}_M(C_x) = \liminf_{k\to\infty} \frac{\sigma_k(x)}{k}.$$

Now, by the Law of Large Numbers we know that for almost every $x \in [0,1]$ we have that

$$\limsup_{k\to\infty} \frac{\sigma_k(x)}{k} = \liminf_{k\to\infty} \frac{\sigma_k(x)}{k} = \frac{1}{2}.$$

Therefore, for almost every $x \in [0,1]$ we also have

$$\underline{\dim}_M(C_x) = \overline{\dim}_M(C_x) = \dim_M(C_x) = 1/2,$$

which completes the proof. $\square$

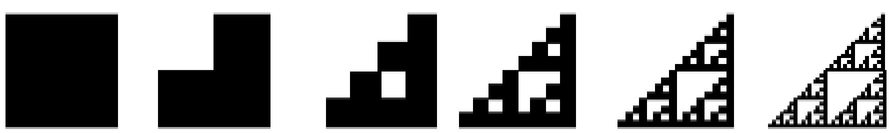


Figure 4.1: The first five iterations of the Sierpiński Gasket

Remark 4.5.4. Note that C_x is an example of a set defined by digit restrictions considered in Chapter 2. Indeed if for $x \in [0,1]$ we let $S_x \subset \mathbb{N}$ be the collection of those i's for which $x_i = 1$ then C_x is in fact the same set as A_{S_x}.

Chapter 5

Bedford-McMullen self-affine sets

5.1 Construction of Bedford-McMullen self-affine sets

In Chapter 2, we mentioned *similarity dimension* where Hausdorff and Minkowski dimensions of a set agreed. Those sets were known as self-similar. Now we introduce self-affine sets which form an important class of sets, which include self-similar sets as a particular case. The two dimensions of this class of self-affine sets usually differ. We would hope to look for a formula for the dimension of self-affine sets that generalizes equation (2.3.1) mentioned in Chapter 2. However, as it turns out no such simple formula exists.

The Hausdorff dimension of a particular class of self-affine sets was computed independently by Tim Bedford and Curt McMullen in the early 1980's, see[2,10]. The main purpose of this chapter is to state and prove their result, which is our main Theorem 5.3.1. We will also calculate Minkowski dimension of these self-affine sets and will see that these two dimensions do not coincide in general.

To construct the Bedford-McMullen self-affine sets we choose integers $n \geq m > 1$. Next we divide the unit square $[0,1]^2$ into $m \times n$ equal closed rectangles, each with width n^{-1} and height m^{-1}. Choose a subset D of these rectangles and throw away the rest. Note, that D may also be thought of as a subset of the set of pairs

$$\{(i,j) : i \in \{0,\ldots,n-1\}, j \in \{0,\ldots,m-1\}.$$

Divide the remaining rectangles into $n \times m$ sub-rectangles each of width n^{-2} and height m^{-2} and keep those corresponding to the same pattern used before. Thus in the second step we have $(\#D)^2$ rectangles. Here and below $\#D$ denotes the cardinality of the set D. We continue this same process and at the kth stage we have a collection of $(\#D)^k$ rectangles of width n^{-k} and height m^{-k}. To attain the next level we subdivide each remaining rectangle into $n \times m$ rectangles as above, and keep the ones corresponding to our pattern. Continuing indefinitely gives us a compact set which we will call the Bedford-McMullen set corresponding to the subset D and will denote by $K(D)$ (see Fig. 5.1) [1].

In Figure 5.1, the first five stages of the construction of $K(D)$ are drown in the case when $n = 3, m = 2$ and

$$D = \{(0,0); (1,1); (2,0)\}.$$

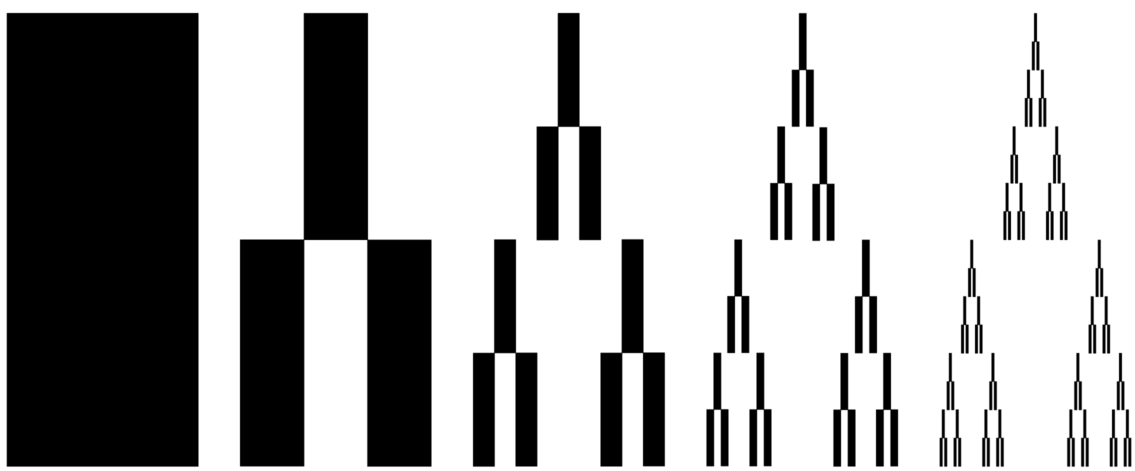


Figure 5.1: The first five stages of the construction of a McMullen set.

5.2 Minkowski dimension of self-affine sets

Theorem 5.2.1. *Suppose every row contains a chosen rectangle. Assume $n > m$. Then*

$$\dim_M(K(D)) = 1 + \log_n \frac{\#D}{m}.$$

Remark 5.2.2. For the proof of this theorem, along with the proof of Theorem 5.3.1 we assume that every row contains a chosen rectangle. In general, if we denote π as the projection onto the second coordinate such that $\#(\pi(D))$ is the number of occupied rows, then

$$\dim_M(K(D)) = \log_m \#(\pi(D)) + \log_n \frac{\#D}{\#(\pi(D))}.$$

Proof. We begin by letting $r = \#D$ be the number of rectangles in our pattern. At the jth stage we have r^j rectangles of width n^{-j} and height m^{-j}. Now let $k = \lfloor \frac{\log n}{\log m} j \rfloor$ be the integer part of $\frac{\log n}{\log m} j$. We can cover each rectangle by M^{k-j} squares of side $m^{-k} (\sim n^{-j})$ and this many squares needed. Now we use our assumption that every row has a rectangle and thus for any generational rectangle R the horizontal projection of $K(D) \cap R$ onto a vertical side of R is the entire side.

Consequently, the total number of squares of side $m^{-k} (\sim n^{-j})$ needed to cover $K(D)$ is $r^j m^{k-j}$ and thus

$$\begin{aligned}
\dim_M(K(D)) &= \lim_{j\to\infty} \frac{\log r^j m^{k-j}}{\log n^j} \\
&= \lim_{j\to\infty} \frac{j \log r + (k-j)\log m}{j \log n} \\
&= \frac{\log r}{\log n} + \frac{\log n}{\log n} - \frac{\log m}{\log n} \\
&= 1 + \log_n(\frac{r}{m}).
\end{aligned}$$

which completes our proof[1]. □

5.3 Hausdorff dimension of self-affine sets

Theorem 5.3.1. *Suppose every row contains a chosen rectangle. Assume $n > m$. Then*

$$\dim_H(K(D)) = \log_m \left(\sum_{j=1}^{m} r(j)^{\log_n m} \right),$$

where $r(j)$ is the number of rectangles of the pattern lying in the j^{th} row[1].

Fix integers $m < n$ and let $\alpha = \frac{\log m}{\log n} < 1$. Now we introduce approximate squares to help us calculate dimension and prove the above theorem.

Definition 5.3.2. Suppose $(x, y) \in [0, 1)^2$ have base n and base m expansions $\{x_k\}$, $\{y_k\}$, respectively. The *approximate square*[1] of generation k at (x, y), $Q_k(x, y)$, is defined to be the closure of the set of points $(x', y') \in [0, 1)^2$ such that the first $\lfloor \alpha k \rfloor$ digits in the base n expansions of x and x' coincide, and the first k digits in the base m expansions of y and y' coincide. We refer to the rectangle $Q_k(x, y)$ as an approximate square of generation k since its width $n^{-\lfloor \alpha k \rfloor}$ and height m^{-k} satisfy:

$$m^{-k} \leq n^{-\lfloor \alpha k \rfloor} \leq nm^{-k}$$

and hence

$$m^{-k} \leq |Q_k(\omega)| \leq (n+1)m^{-k}.$$

Proof of Theorem 5.3.1. Any probability vector $\{p(d) : d \in D\} = \mathbf{p}$ defines a probability measure $\mu_{\mathbf{p}}$ on $K(D)$ which is the image of the product measure $\mathbf{p}^{\mathbb{N}}$ under the representation map

$$R : D^{\mathbb{N}} \to K(D) \tag{5.3.1}$$

given by

$$\{(a_k, b_k)\}_{k=1}^{\infty} \to \sum_{k=1}^{\infty} (a_k n^{-k}, b_k m^{-k}).$$

Any such measure is supported on $K(D)$, so the dimensions of these measures all give lower bounds for the dimension of $K(D)$. We shall show that the supremum of these dimensions is exactly $\dim_H(K(D))$. In fact, we will restrict our attention to measures coming from probability vectors $\mathbf{p}$ such that

$$p(d) \text{ depends only on the second coordinate of } d, \tag{5.3.2}$$

i.e. all rectangles in the same row get the same mass.

Let (x, y) be in $K(D)$. Suppose $\{x_\nu\}$, $\{y_\nu\}$ are the n-ary and m-ary expansions of x and y. We claim that

$$\mu_{\mathbf{p}}(Q_k(x,y)) = \prod_{\nu=1}^{k} p(x_\nu, y_\nu) \prod_{\alpha k+1}^{k} r(y_\nu), \tag{5.3.3}$$

where for $d = (i, j) \in D$ we denote $r(d) = r(j)$, the number of elements in row j. To see this, note that the $n^{-k} \times m^{-k}$ rectangle defined by specifying the first k digits in the base n expansion of x and the first k digits in the base m expansion of y have $\mu_{\mathbf{p}}$-measure $\prod_{\nu=1}^{k} p(x_\nu, y_\nu)$. The approximate square $Q_k(x, y)$ contains

$$r(y_{\alpha k+1}) \cdot r(y_{\alpha k+2}) \cdot \dots \cdot r(y_k)$$

such rectangles, all with the same $\mu_{\mathbf{p}}$-measure by our assumption (5.3.2), and so (5.3.3) follows. Now take logarithms in (5.3.3),

$$\log(\mu_{\mathbf{p}}(Q_k(x,y))) = \sum_{\nu=1}^{k} \log p(x_\nu, y_\nu) + \sum_{\alpha k+1}^{k} \log r(y_\nu). \tag{5.3.4}$$

Since $\{(x_\nu, y_\nu)\}_{\nu \geq 1}$ are i.i.d (independent and identically distributed) random variables with respect to $\mu_{\mathbf{p}}$, the Law of Large Numbers yields for $\mu_{\mathbf{p}}$-a.e. (x, y) :

$$\lim_{k\to\infty} \frac{1}{k} \log(\mu_{\mathbf{p}}(Q_k(x,y))) = \sum_{d\in D} p(d) \log p(d) + (1-\alpha) \sum_{d\in D} p(d) \log r(d). \tag{5.3.5}$$

The proof of Billingsley's lemma 4.1.3 extends to this setting and implies

$$\dim(\mu_{\mathbf{p}}) = \frac{1}{\log m} \sum_{d\in D} p(d) \left(\log \frac{1}{p(d)} + \log(r(d)^{\alpha-1}) \right). \tag{5.3.6}$$

An easy and well known calculation says that if $\{a_k\}_{k=1}^{n}$ are real numbers, then the maximum of the function

$$F(\mathbf{p}) = \sum_{k=1}^{n} p_k \log \frac{1}{p_k} + \sum_{k=1}^{n} p_k a_k$$

over all probability measure $\mathbf{p}$ is attained at $p_k = e^{a_k}/\sum_\ell e^{a_\ell}$, $k = 1, ..., n$. This is known as Boltzmann's principle[1], which is proved in Appendix B. In the case at hand, it says that $\dim(\mu_{\mathbf{p}})$ will be maximized if

$$p(d) = \frac{1}{Z} r(d)^{\alpha-1} \tag{5.3.7}$$

where

$$Z = \sum_{d\in D} r(d)^{\alpha-1} = \sum_{j=0}^{m-1} r(j)^{\alpha}.$$

For the rest of the proof, we fix this choice of $\mathbf{p}$ and write μ for $\mu_{\mathbf{p}}$. Note that

$$\dim(\mu) = \log_m(Z),$$

so this is a lower bound for $\dim_H(K(D))$. To obtain an upper bound, denote

$$S_k(x,y) = \sum_{\nu=1}^{k} \log r(y_\nu).$$

Note that $\frac{1}{k} S_k(x,y)$ is uniformly bounded. Using (5.3.7), rewrite (5.3.4) in the form

$$\begin{aligned}
\log \mu(Q_k(x,y)) &= \sum_{\nu=1}^{k} \log \frac{1}{Z} r(y_\nu)^{\alpha-1} + \left(\sum_{\nu=1}^{k} \log r(y_\nu) - \sum_{\nu=1}^{\alpha k} \log r(y_\nu) \right) \\
&= -\sum_{\nu=1}^{k} \log Z + (\alpha - 1) S_k(x,y) + S_k(x,y) - S_{\alpha k}(x,y).
\end{aligned}$$

Thus

$$\log \mu(Q_k(x,y)) + k \log Z = \alpha S_k(x,y) - S_{\alpha k}(x,y).$$

Therefore,

$$\frac{1}{\alpha k}\log \mu(Q_k(x,y)) + \frac{1}{\alpha}\log Z = \frac{S_k(x,y)}{k} - \frac{S_{\alpha k}(x,y)}{\alpha k}. \tag{5.3.8}$$

Summing the right hand side of (5.3.8) along $k = \alpha^{-1}, \alpha^{-2}, \ldots$ gives a telescoping series[15].

Since $\frac{S_k(x,y)}{k}$ remains bounded for all k, it is easy to see

$$\limsup_{k\to\infty}\left(\frac{S_k(x,y)}{k} - \frac{S_{\alpha k}(x,y)}{\alpha k}\right) \geq 0,$$

since otherwise the sum would tend to $-\infty$. Thus, by (5.3.8) for every $(x,y) \in K(D)$ we have

$$\limsup_{k\to\infty}(\log \mu(Q_k(x,y)) + k\log Z) \geq 0.$$

This implies

$$\liminf_{k\to\infty}\frac{\log \mu(Q_k(x,y))}{-k} \leq \log Z.$$

Since $m^{-k} \leq |Q_k(x,y)| \leq (n+1)m^{-k}$, the last inequality, along with Billingsley's lemma implies that

$$\dim_H(K(D)) \leq \log_m(Z).$$

Combining this with the lower bound above, we get

$$\dim_H(K(D)) = \frac{\log Z}{\log m} = \log_m Z$$

which is exactly what we wanted to prove, Theorem 5.3.1. $\square$

Chapter 6

Conclusion

One way of comparing fractals is by studying their fractal dimension. This means studying small sets and how well they can be covered by balls (squares, triangles, etc.). We introduced the concept of Minkowski dimension, explored its basic properties and calculated it for different sets. Next, we defined Hausdorff dimension, established some of its basic properties and studied several important examples. Then we introduced the Mass Distribution Principle in order to obtain lower bounds for Hausdorff dimension. We looked at the dimension of a measure and established a refined version of the Mass Distribution Principle, known as Billingsley's lemma. We used these techniques to study an interesting example of the dimension of random vertical slices of the Sierpiński gasket and of other self-similar sets. We concluded with the construction of the Bedford-McMullen self-affine sets and the proof of the formula for the Hausdorff dimension of these type of sets.

Fractals are very beautiful objects that are found everywhere in nature. They have been used to model a variety of problems in many different scientific and non-scientific fields, most notably in dynamical systems. Benoit Mandelbrot is known as the "father of fractals" for all of his contributions to the field. As mentioned before, there is still much to be studied and discovered when it comes to fractals.

Below we have a table that has the dimension values of various sets mentioned and defined in this report. The table includes both the Minkowski and Hausdorff dimensions of

Set	Minkowski Dimension	Hausdorff Dimension
A point $\{p\}$	$\dim_M(\{p\}) = 0$	$\dim_H(\{p\}) = 0$
$[0,1]$	$\dim_M([0,1]) = 1$	$\dim_H([0,1]) = 1$
$[0,1]^N$	$\dim_M([0,1]^N) = N$	$\dim_H([0,1]^N) = N$
$\mathbb{Q} \cap [0,1]$	$\dim_M(\mathbb{Q} \cap [0,1]) = 1$	$\dim_H(\mathbb{Q} \cap [0,1]) = 0$
$(\{0\} \bigcup_{n=1}^{\infty} \{\frac{1}{n}\})$	$\dim_M(\{0\} \bigcup_{n=1}^{\infty} \{\frac{1}{n}\}) = \frac{1}{2}$	$\dim_H(\{0\} \bigcup_{n=1}^{\infty} \{\frac{1}{n}\}) = 0$
Middle third Cantor set : $\mathcal{C}$	$\dim_M(\mathcal{C}) = \log_3 2$	$\dim_H(\mathcal{C}) = \log_3 2$
Koch Snowflake : K	$\dim_M(K) = \log_3 4$	$\dim_H(K) = \log_3 4$
Sierpiński Carpet : SC	$\dim_M(SC) = \log_3 8$	$\dim_H(SC) = \log_3 8$
Sierpiński Gasket : SG	$\dim_M(SG) = \log_2 3$	$\dim_H(SG) = \log_2 3$
A random vertical : VG slice of the gasket	$\dim_M(VG) = \frac{1}{2}$	$\dim_H(VG) = \frac{1}{2}$
Bedford-McMullen Set : BM	$\dim_M(BM) = 1 + \frac{\log \frac{3}{2}}{\log 3}$	$\dim_H(BM) = \frac{\log(1+2^{\log_3 2})}{\log 2}$

Table 6.1: Dimension values of various sets.

the specified sets.

Appendix A

Law of Large Numbers:An important technique in fractal analysis

Definition A.0.3. We call $\mathbf{P}$ a *probability* or *probability measure*[4] if $\mathbf{P}$ assigns a number $\mathbf{P}(A)$ to each A in $\mathcal{F}$. Here $\mathcal{F}$ is the events, i.e. a collection of subsets $\mathcal{F} = \{F_i\}_{i\in I}$ where $F_i \subseteq \Omega$ for every $i \in I$. $\mathbf{P}(A)$ is the probability that the event A will occur and Ω is the set of all possible outcomes of an experiment called the *sample space*, such that the following conditions hold true:

1. $0 \leq \mathbf{P}(A) \leq 1$ for all $A \in \mathcal{F}$,
2. $\mathbf{P}(\emptyset) = 0$ and $\mathbf{P}(\Omega) = 1$,
3. and, if $A_1, A_2, ...$ are disjoint events in $\mathcal{F}$, then

$$\mathbf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbf{P}(A_i).$$

Probability is in fact a measure: $\mathcal{F}$ is the σ-algebra on which $\mathbf{P}$ is defined. The only additional property is that the measure of the total space is 1, i.e. $\mathbf{P}(\Omega) = 1$, as briefly mentioned in the previous chapter[4].

Definition A.0.4. The triple $(\Omega, \mathcal{F}, \mathbf{P})$ is called a *probability space*[4] if $\mathcal{F}$ is an event space of subsets of Ω, i.e. $\mathcal{F}$ is a σ-algebra and $\mathbf{P}$ is a probability measure defined on the sets of $\mathcal{F}$.

Theorem A.0.5. (Law of Large Numbers): *Let* $\{f_n\}$, *with* $n = 1, 2, \ldots$ *be a sequence of orthogonal functions on a probability space* $(\mathrm{X}, d\nu)$ *and suppose* $E({f_i}^2) = \int |f_i|^2 d\nu \leq 1$. *Then*

$$\frac{1}{n} S_n = \frac{1}{n} \sum_{k=1}^{n} f_k \to 0,$$

almost everywhere, with respect to ν *(denoted* ν*-a.e.), as* $n \to \infty$.

Proof. First, let us assume that if $\{g_n\}$ is a sequence of functions on a probability space $(X, d\nu)$ such that

$$\sum_n \int |g_n|^2 d\nu < \infty,$$

then $\sum_n |g_n|^2 < \infty$, ν-a.e. and thus $g_n \to 0$ ν-a.e. We can use this assumption to prove the Law of Large Numbers for $n \to \infty$ along the sequence of squares. That is to say,

$$\begin{aligned}
\int \left(\frac{1}{n} S_n \right)^2 d\nu &= \frac{1}{n^2} \int (S_n)^2 \, d\nu = \frac{1}{n^2} \int \left(\sum_{i=1}^{n} f_i \right)^2 d\nu \text{ (by our assumption)} \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \int f_i^2 d\nu \text{ (by orthogonality)} \\
&\leq \left(\frac{1}{n^2} \right) n = \frac{1}{n}.
\end{aligned}$$

Let $\frac{1}{n} S_n = h_n$, then if we set $g_n = h_{n^2}$, from our assumption, we have

$$\begin{aligned}
\sum_{n=1}^{\infty} \int |g_n|^2 d\nu &= \sum_{n=1}^{\infty} \int \left(\frac{1}{n^2} |S_{n^2}| \right)^2 d\nu = \sum_{n=1}^{\infty} \frac{1}{n^4} \int |S_{n^2}|^2 d\nu \text{ (by substitution)} \\
&\leq \sum_{i=1}^{\infty} \left(\frac{1}{n^4} \right) n^2 \text{ (by orthogonality)} \\
&= \sum_{i=1}^{\infty} \frac{1}{n^2} < \infty.
\end{aligned}$$

Therefore, since the right-hand side of the above inequality is summable, our observation above implies $g_n \to 0$ ν-a.e. Identically, we have $\frac{1}{n^2} S_{n^2} \to 0$ almost everywhere. Now, we want to show $h_n \to 0$.

To handle the limit over all integers, we choose an n such that

$$m^2 \leq n < (m+1)^2 \Longleftrightarrow m \leq \sqrt{n} < m+1$$

and set $m(n) = \lfloor \sqrt{n} \rfloor$ (the integer part of n), the largest integer not greater than n. Thus, we have

$$\begin{aligned}
\int |\frac{1}{m^2} S_n - \frac{1}{m^2} S_{m^2}|^2 d\nu &= \frac{1}{m^4} \int |S_n - S_{m^2}|^2 d\nu = \frac{1}{m^4} \int | \left(\sum_{i=m^2+1}^{n} f_i \right) |^2 d\nu \\
&= \frac{1}{m^4} \int \sum_{i=m^2+1}^{n} |f_i|^2 d\nu \leq \frac{1}{m^4} \sum_{i=m^2+1}^{n} \int |f_i|^2 d\nu \\
&\leq \frac{1}{m^4}(n - (m^2+1)) \leq \frac{1}{m^4}((m+1)^2 - (m^2+1)) \\
&\leq \frac{2}{m^3},
\end{aligned}$$

since the sum has at most $2m$ terms, each of size at most 1. Set

$$g_n = \frac{S_n}{m(n)^2} - \frac{S_{m(n)^2}}{m(n)^2}.$$

Hence, since $m = m(n)$ is associated to at most $2m+1$ different n's we get

$$\begin{aligned}
\sum_{n=1}^{\infty} \int |g_n|^2 d\nu &\leq \sum_{n=1}^{\infty} \frac{2}{m(n)^3} \\
&\leq \sum_{m} (2m+1) \frac{2}{m^3} \\
&< \infty.
\end{aligned}$$

Therefore, by our original assumption, $g_n \to 0$ almost everywhere,

$$\implies \quad \frac{1}{m(n)^2} S_n \to 0 \ \ \nu\text{-a.e.},$$
$$\implies \quad \frac{1}{n} S_n \to 0 \ \ \nu\text{-a.e.}.$$

Consequently, we have proved that

$$\frac{1}{n} S_n = \sum_{k=1}^{n} f_k \to 0$$

almost everywhere, with respect to ν, as $n \to \infty$ [1]. □

Appendix B

Boltzmann's Principle

In order to prove Boltzmann's Principle, we must first introduce a concept known as the Method of Lagrange Multipliers.

Suppose that we want to maximize or minimize a function of n variables

$$f(x) = f(x_1, x_2, \ldots, x_n) \text{ for } x = (x_1, x_2, \ldots, x_n) \tag{B.0.1}$$

subject to p constraints

$$g_1(x) = c_1, \;\; g_2(x) = c_2, \;\; \ldots, \;\; \text{and } \; g_p(x) = c_p \tag{B.0.2}$$

Theorem B.0.6. (Lagrange's Theorem): *Assuming appropriate smoothness conditions, minimum or maximum of $f(x)$ subject to the constraints (B.0.2) that is not on the boundary of the region where $f(x)$ and $g_j(x)$ are defined can be found by introducing p new parameters $\lambda_1, \lambda_2, \ldots, \lambda_p$ and solving the system*

$$\frac{\partial}{\partial x_i}\left(f(x) + \sum_{j=1}^{p} \lambda_j g_j(x)\right) = 0, \quad 1 \leq i \leq n \tag{B.0.3}$$

$$g_j(x) = c_j, \quad 1 \leq j \leq p. \tag{B.0.4}$$

This amounts to solving $n + p$ equations for the $n + p$ real variables in x and λ. The proof can be found in any multivariate calculus book. Now we state and prove Boltzmann's Principle, using Theorem (B.0.6) in the proof.

Lemma B.0.7. (BOLTZMANN'S PRINCIPLE): *If $\{a_k\}_{k=1}^n$ are real then the max of the function*

$$F(p) = \sum p_k \log \frac{1}{p_k} + \sum p_k a_k \tag{B.0.5}$$

over all probability vectors p is attained at $p_k = \frac{e^{a_k}}{\sum_\ell e^{a_\ell}}$, where $k = 1, \ldots, n$.

Proof. We want to maximize $F(p_1, p_2, \ldots, p_n)$ subject to the constraint $p_1+p_2+\ldots+p_n = 1$. This is done using the method of Lagrange multipliers. For that, let $g(p_1, \ldots, p_n) = p_1 + \ldots + p_n - 1$. At the maximum, we will have $\nabla F = -\lambda \nabla g$. Therefore, we calculate

$$\begin{aligned} \frac{\partial F}{\partial p_i} &= \frac{\partial}{\partial p_i}\left(p_i \log \frac{1}{p_i} + p_i a_i\right) \\ &= \log \frac{1}{p_i} + p_i\left(-\frac{1}{p_i}\right) + a_i \\ &= -\log p_i - 1 + a_i \end{aligned}$$

Since $\frac{\partial g}{\partial p_i} = 1$, we obtain

$$\begin{aligned} -\log p_i - 1 + a_i &= -\lambda \\ \implies \log p_i &= \lambda - 1 + a_i \\ \implies p_i &= e^{\lambda-1} \cdot e^{a_i}. \end{aligned}$$

Since

$$\begin{aligned} \sum_{i=1}^n p_i &= \sum_{i=1}^n e^{\lambda-1} e^{a_i} \\ &= e^{\lambda-1} \sum_{i=1}^n e^{a_i} \\ &= 1 \end{aligned}$$

we obtain

$$e^{\lambda-1} = \frac{1}{\sum_{i=1}^{n} e^{a_i}}$$

and therefore

$$p_i = \frac{e^{a_i}}{\sum_{i=1}^{n} e^{a_i}}$$

as desired.

$\square$

Bibliography

[1] Bishop, Christopher J. and Yuval Peres, *Fractal Sets in Probability and Analysis*, soon to be published.

[2] Bedford, T., *Crinkly curves, Markov partitions and dimension.* PhD thesis, University of Warwick, 1984.

[3] Devaney, Robert L., *A First Course in Chaotic Dynamical Systems:Theory and Experiment*, Perseus Books Publishing, L.L.C. (1992).

[4] Falconer, Kenneth J., *Fractal Geometry: Mathematical Foundations and Applications*, second edition, John Wiley and Sons Ltd. (2003).

[5] Falconer, Kenneth J., *The Geometry of Fractal Sets*, Cambridge University Press (1985).

[6] Falconer, Kenneth J., *Techniques in Fractal Geometry*, John Wiley and Sons Ltd. (1997).

[7] Leifsson, Patrik, *Fractal sets and dimensions*, Linköping University Electronic Press (2006).

[8] Mandelbrot, Benoit B., *The Fractal Geometry of Nature*, W. H. Freeman and Company (1983).

[9] Mattila, Pertti, *Geometry of Sets and Measures in Euclidean Spaces: fractals and rectifiability*, Cambridge University Press (1995).

[10] McMullen, Curt *The Hausdorff dimension of general Sierpiński carpets*, Nagoya Math. J. 96 (1984), 1–9.

[11] Moran, Patrick A.P., *Additive functions of intervals and Hausdorff measure*, Proc. Cambridge Philos. Soc. (1946).

[12] Mörters, Peter and Yuval Peres, *Browian Motion: Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press (2010).

[13] Pesin, Yakov B., *Dimenison Theory in Dynamical Systems: Contemporary Views and Applications*, The University of Chicago Press (1997).

[14] Pollicott, Mark, *Lectures on Fractals and Dimension Theory*, The University of Warwick (2011).

[15] Rudin, Walter, *Principles of Mathematical Analysis*, third edition, McGraw-Hill, Inc. (1964).

[16] Seal, David C., *An Introduction to Fractals and Hausdorff Measures*, University of Utah (2011).

[17] Shane, Christopher, *An Exploration of Hausdorff Dimension*, Heckman Bindery, Inc.(2000).

[18] Strogatz, Steven H., *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*, Perseus Books Publishing, L.L.C. (1994).

[19] Wheeden, Richard L. and Antoni Zygmund, *Measure and Integral: An Introduction to Real Analysis*, Marcel Dekker, Inc. (1977).