Introduction to fractal dimension

by

Ghder Aburamyah

B.S., Almajmaah University, Saudi Arabia, 2011

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

2018

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 40 years, this small branch of mathematics has developed extensively. Fractals can be defined as those sets which have non-integer Hausdorff or Minkowski dimension. In this report, we introduce certain definitions of fractal dimensions, which can be used to measure a set's fractal degree. We introduce Minkowski dimension and Hausdorff dimension and explore some examples where they coincide, as well as other examples where they do not.

Table of Contents

Acknowledgments

All praises to Allah for the strength and his blessing in completing this report. All thanks to Allah for giving me the patience to finally be done after all difficulties and challenges. With great appreciation, I would like to thank my major advisor Dr. Hrant Hakobyan for his help and guidance during the preparation of this report. I also would like to express my gratitude to the members of my examination committee, Dr. Sarah Reznikoff and Dr. Anna Zemlyanova. A special thanks to my family. Words can not express how grateful I am to my mother and father for all the love, support and inspiration. Your prayer for me was what sustained me thus far. Thank you to my sisters and brothers, they kept encouraging me towards success. I would also like to thank my beloved husband for supporting me for everything, and especially I can't thank you enough for encouraging me throughout this experience. To my two lovely children, Majid and Abdulrahman who are the pride and joy of my life. I love you more than anything and I appreciate all your patience and support during mommy's Master studies. Last but not least I would like to thank Saudi Ministry of Education for their great scholarship and the opportunity to study abroad.

Chapter 1

Introduction

Everyone knows that the dimension of a line, a square, and a cube are one, two, and three, respectively. Also, we can measure the length of a line, the area of a square, and volume of a cube. Though, if we need to measure the dimension of the brain or lungs, then we need the notion of fractal dimension. Fractal dimension is very important in mathematics because it allows us to measure the complexity of non-smooth objects. The notion of dimension is central to fractal geometry. Roughly, dimension indicates how much space a set occupies near each of its points. In this report, we look at the two most common types of fractal dimensions, which are known as Minkowski dimension and Hausdorff dimension. Then we explore 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.

Most of the material in this report is quite standard for geometric measure theory and can be found in the well known books by Falconer[?] and Bishop and Peres[?].

Chapter 2

Minkowski and Box Counting Dimensions

In this section we will discuss a type of Fractal dimension called Minkowski dimension. Minkowski dimension is also 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 easier to calculate or estimate.

2.1 Definitions

Below we often consider subsets of Euclidean space R^n even though most of the definitions work in a general metric space X. The *distance* between points $x = (x_1, x_2, ..., x_n) \in R^n$ and $y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$ will be denoted by

$$|x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}. \tag{2.1}$$

Recall that a metric d on X is a function $d : X \times X \to R$ such that for all $x, y, z \in X$:

i) $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only if $x = y$;

ii) $d(x, y) = d(y, x)$ (symmetry);

iii) $d(x, y) \leq d(x, z) + d(z, x)$ (tringle inequality).

A metric space (X, d) is a set X with a metric d defined on X.

If (X, d) is a metric space, $y \in X$, and radius $r > 0$ then the *Open Ball and Closed Ball* centered at y with radius r are defined, respectively, as follows:

$$B(y, r) = \{x \in X : d(x, y) < r\}. \tag{2.2}$$

$$\overline{B}(y, r) = \{x \in X : d(x, y) \leq r\}. \tag{2.3}$$

Recall also that the diameter of a set E in a metric space X is defined as follows:

$$\operatorname{diam}(E) = \sup\{d(x, y) \,|\, x, y \in E\}. \tag{2.4}$$

We say E is a *totally bounded* set in a metric space X if for any $\epsilon > 0$, it can be covered by a finite number of sets of diameter ϵ. This means that for every $\epsilon > 0$ there is a number $M = M(\epsilon) \in \mathbb{N}$ and sets $U_i \subset X, i = 1, \ldots, M$, such that $E \subset \cup_{i=1}^{M} E_i$.

Definition 2.1.1 (Minkowski Dimension)**.** *Let (X, d) be a metric space and $E \subset X$ be a totally bounded set. The upper and lower Minkowski dimensions of E are defined, respectively, as follows:*

$$\overline{\dim}_M(E) = \limsup_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}}$$
$$\underline{\dim}_M(E) = \liminf_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}},$$

where $N(E, \epsilon)$ is the smallest number of sets with diameter ϵ needed to cover E.

If $\overline{\dim}_M(E) = \underline{\dim}_M(E)$, then the common value is called the Minkowski dimension of E and we write

$$\dim_M(E) = \lim_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}}.$$

Remark 2.1.2. *From the definition it follows that to estimate $\dim_M E$ (if it exists) from above, one only needs to take an arbitrary covering of E by sets of diameter ϵ. Indeed, if*

$E \subset \bigcup_{i=1}^{N} E_i$ *where* $E_i \subset X$*, then* $N(E, \epsilon) \leq N$*. Lower bounds are harder to obtain, because to estimate* $N(E, \epsilon)$ *from below we need to consider all the covering and find the "optimal" one (i.e. the one with fewest possible members of the covering) for each* $\epsilon > 0$*.*

Remark 2.1.3. *If the Minkowski dimension exists and is equal to d, then it follows from the definition that the smallest number of sets with diameter* ϵ *needed to cover* E *grows like* $(\frac{1}{\epsilon})^{d+o(1)} \to \infty$ *as* $\epsilon \to 0$*.*

Remark 2.1.4. *It follows immediately from the definition that*

$$\dim_M E = \dim_M(\overline{E}) \tag{2.5}$$

where $\overline{E}$ *is the closure of* E*.*

By using coverings by balls instead of arbitrary coverings we can define a similar concept as follows.

Definition 2.1.5 (Box-counting dimension)**.** *Let* E *be a bounded set in a metric space and let* $N_B(E, \epsilon)$ *be the smallest number of closed balls of radius* ϵ *required to cover the set* E*. The limits*

$$\overline{\dim}_B(E) = \limsup_{\epsilon \to 0} \frac{\log N_B(E, \epsilon)}{\log \frac{1}{\epsilon}}$$
$$\underline{\dim}_B(E) = \liminf_{\epsilon \to 0} \frac{\log N_B(E, \epsilon)}{\log \frac{1}{\epsilon}},$$

are called the upper and lower box-counting dimensions of E*, respectively. If the limit*

$$\dim_B(E) = \lim_{\epsilon \to 0} \frac{\log N_B(E, \epsilon)}{\log \frac{1}{\epsilon}}.$$

exists, then we call it the box-counting dimension of E*.*

Lemma 2.1.6. *For every bounded set E in a metic space we have*

$$\dim_B(E) = \dim_M(E).$$

Proof. First, observe that

$$N_B(E, \epsilon) \leq N(E, \epsilon) \leq N_B(E, \frac{\epsilon}{2}) \tag{2.6}$$

Strict inequalities can occur in general metric space. Indeed, the first inequality is true because for every set there is a ball with the same diameter that contains that set. So every ball can contain several sets with the same diameter. Therefore we have

$$\dim_B E = \lim_{\epsilon \to 0} \frac{\log N_B(E, \epsilon)}{\log \frac{1}{\epsilon}} \leq \lim_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}} = \dim_M E.$$

On the other hand, $N(E, \epsilon) \leq N_B(E, \frac{\epsilon}{2})$, for $\dim_M E \leq \dim_B E$ we use the second inequality (2.6),

$$\begin{aligned}
\dim_M E = \lim_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}} &\leq \lim_{\epsilon \to 0} \frac{\log N_B(E, \frac{\epsilon}{2})}{\log \frac{1}{\epsilon} \cdot \frac{2}{2}} \\
&= \lim_{\epsilon \to 0} \frac{\log N_B(E, \frac{\epsilon}{2})}{\log \frac{2}{\epsilon} + \log \frac{1}{2}} \\
&= \lim_{\epsilon \to 0} \frac{\log N_B(E, \frac{\epsilon}{2})}{\log \frac{2}{\epsilon}[1 + \frac{\log \frac{1}{2}}{\log \frac{2}{\epsilon}}]} \\
&= \lim_{\epsilon \to 0} \frac{\log N_B(E, \frac{\epsilon}{2})}{\log \frac{2}{\epsilon}} \\
&= \dim_B M,
\end{aligned}$$

and the proof is complete. $\square$

The following remark will be quite useful in estimating Minkowski dimension of many sets.

Remark 2.1.7. *If X is a set such that if $x, y \in X$ the dist$(x, y) \geq \epsilon$, then we say that X is $\epsilon -$ separated. Let $N_{sep}(C, \epsilon)$ be the number of elements in a maximal $\epsilon -$ separated subset X of C. Obviously, any set of diameter $\frac{\epsilon}{2}$ contains at most one point of an $\epsilon -$ separated set X, so $N_{sep}(C, \epsilon) \leq N(C, \frac{\epsilon}{2})$. Conversely, every point of a set C is within ϵ of a maximal $\epsilon -$ separated subset X. Thus $N(C, \epsilon) \leq N_{sep}(C, \epsilon)$. Therefore, if we replace $N(C, \epsilon)$ with $N_{sep}(C, \epsilon)$ then the upper and lower Minkowski dimensions will give us the same values, just as in the proof of Lemma 2.1.6.*

2.2 Minkowski dimension of finite sets

Lemma 2.2.1. *If E is a finite set then* $\dim_M(E) = 0$.

Proof. Let E be a finite set $\{a_i\}_{i=1}^n$ in a metric space (X, d_X). Clearly we have $N(E, \epsilon) \leq n$ for all $\epsilon > 0$. On the other hand, if

$$\delta := \min_{i,j} d_X(a_i, a_j)$$

and if $\epsilon < \delta/2$ then we can see that $N(E, \epsilon) \geq n$. Indeed, if we have two points whose distance is δ then they cannot belong to a single ball of diameter δ or equivalently of radius $\epsilon < \delta/2$. So if the diameter of a ball is small enough then every single ball can have at most one point. Therefore, for $\epsilon < \delta/2$ we have $N(E, \epsilon) = n$ and

$$\begin{aligned} \dim_M(E) &= \lim_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}} \\ &= \lim_{\epsilon \to 0} \frac{\log n}{\log \frac{1}{\epsilon}} = 0. \end{aligned}$$

$\square$

2.3 The interval

Lemma 2.3.1. *If* $E = [0, 1]$ *then* $\dim_M E = 1$.

Proof. Let $E = [0, 1]$. We need at least $\lceil \frac{1}{\epsilon} \rceil$ intervals of length ϵ to cover E. Since $\lceil \frac{1}{\epsilon} \rceil < \frac{1}{\epsilon} + 1$, i.e. $N(E, \epsilon) \leq \frac{1}{\epsilon} + 1$, we obtain

$$\begin{aligned}
\overline{\dim}_M(E) = \overline{\lim}_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}} &\leq \lim_{\epsilon \to 0} \frac{\log(\frac{1}{\epsilon} + 1)}{\log(\frac{1}{\epsilon})} = \\
&= \lim_{\epsilon \to 0} \frac{\log(\frac{1+\epsilon}{\epsilon})}{\log(\frac{1}{\epsilon})} \\
&= \lim_{\epsilon \to 0} \frac{\log(1+\epsilon) - \log(\epsilon)}{-\log(\epsilon)} \\
&= 1 - \lim_{\epsilon \to 0} \frac{\log(1+\epsilon)}{\log(\epsilon)} = 1.
\end{aligned}$$

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

Similarly, using the fact that $N(E, \epsilon) \geq \frac{1}{\epsilon} - 1$, we obtain

$$\begin{aligned}
\underline{\dim}_M(E) = \underline{\lim}_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}} &\geq \lim_{\epsilon \to 0} \frac{\log(\frac{1}{\epsilon} - 1)}{\log(\frac{1}{\epsilon})} = \\
= \lim_{\epsilon \to 0} \frac{\log(\frac{1-\epsilon}{\epsilon})}{\log(\frac{1}{\epsilon})} \\
= \lim_{\epsilon \to 0} \frac{\log(1-\epsilon) - \log(\epsilon)}{-\log(\epsilon)} \\
= 1 - \lim_{\epsilon \to 0} \frac{\log(1-\epsilon)}{\log(\epsilon)} = 1.
\end{aligned}$$

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

$$\dim_M(E) = \lim_{\epsilon \to 0} \frac{\log N([0, 1], \epsilon)}{\log \frac{1}{\epsilon}} = 1.$$

$\square$

2.4 The Middle-Thirds Cantor Set

Consider the following collection of intervals,

$$\begin{aligned}
C_0 &= [0,1]\,,\\
C_1 &= \left[0,\frac{1}{3}\right] \cup \left[\frac{2}{3},1\right],\\
C_2 &= \left[0,\frac{1}{9}\right] \cup \left[\frac{2}{9},\frac{3}{9}\right] \cup \left[\frac{6}{9},\frac{7}{9}\right] \cup \left[\frac{8}{9},1\right],\\
&\ldots\ldots\ldots\ldots\ldots
\end{aligned}$$

Thus, at every step the new intervals are obtained by removing at each step the middle-third of every interval remaining from the previous step. By induction, C_n is a union of 2^n closed intervals $I_{n,1},\ldots,I_{n,2^n}$ and the middle-thirds Cantor set is defined as the following infinite intersection,

$$C = \bigcap_{n=0}^{\infty} C_n = \bigcap_{n=0}^{\infty}\bigcup_{k=1}^{2^n} I_{n,k},$$

Where $C_n = \bigcup\limits_{k=1}^{2^n} I_{n,k}$. As an intersection of closed sets C is closed.

Lemma 2.4.1. *If C is the middle-third cantor set then*

$$\dim_M C = \log_3 2.$$

Proof. By construction, C has a covering with 2^n intervals of length $\frac{1}{3^n}$. Therefore,

$$\begin{aligned}
N(C,\frac{1}{3}) &\le 2,\\
N(C,\frac{1}{3^2}) &\le 2^2,\\
&\ldots..\\
N(C,\frac{1}{3^n}) &\le 2^n
\end{aligned}$$

Fix $\epsilon > 0$ and choose n so that $3^{-n} \le \epsilon < 3^{-n+1}$. From the construction it follows that for $\epsilon \ge 3^{-n}$ we have $N(C, \epsilon) \le 2^n$. Therefore,

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

Conversely, for $\epsilon < 3^{-n+1}$ we have $N_{sep}(C, \epsilon) \ge 2^n$ see Remark 2.1.7, hence

$$\underline{\dim}_M(C) \ge \frac{\log 2}{\log 3} = \log_3 2$$

Since the upper and lower Minkowski dimensions are equal, the Minkowski dimension exists, and $\dim_M(C) = \log_3 2$. $\square$

2.5 Generalized Cantor sets

Let $0 < \alpha < 1$. Let C_α be the Cantor set obtained by removing the middle $\alpha' th$ part at each stage. Namely, at every step the new intervals are obtained by removing from every interval I remaining from the previous step an interval of length $\alpha|I|$ from the middle of I.

Lemma 2.5.1. *If C_α is the Cantor set corresponding to $\alpha \in (0,1)$ as above, then*

$$\dim_M(C_\alpha) = \frac{\log 2}{\log 2 + \log \frac{1}{1-\alpha}}.$$

Proof. Note that 2^n intervals of length $(\frac{1-\alpha}{2})^n$ cover C_α, by construction. That means if $\epsilon \ge (\frac{1-\alpha}{2})^n$, then 2^n intervals of length ϵ will cover C_α. Therefore, $N(C, \epsilon) \le 2^n$.
So, $N(C, \epsilon) \le 2^n$ for $\epsilon \ge (\frac{1-\alpha}{2})^n$ and we can estimate the upper Minkowski dimension, as

follows

$$\begin{aligned}
\overline{\dim}_M(C) &\leq \limsup_{\epsilon\to 0} \frac{\log N(C,\epsilon)}{\log\frac{1}{\epsilon}} \\
&= \limsup_{n\to 0} \frac{\log 2^n}{\log(\frac{2}{1-\alpha})^n} \\
&= \limsup_{n\to 0} \frac{n\log 2}{n\log\frac{2}{1-\alpha}} \\
&= \frac{\log 2}{\log 2 + \log\frac{1}{1-\alpha}}.
\end{aligned}$$

On the other hand, if we have $\epsilon < (\frac{1-\alpha}{2})^n$, then each interval of diameter ϵ will not be able to cover an interval of generation n. So, we need at least 2^n intervals of length ϵ to cover the interval. Therefore, $N_{sep}(C,\epsilon) \geq 2^n$, and the lower Minkowski dimension of E can be estimated as follows

$$\underline{\dim}_M(C) \geq \frac{\log 2}{\log 2 + \log\frac{1}{1-\alpha}}.$$

Therefore, the Minkowski dimension exists and is equal to $\frac{\log 2}{\log 2+\log\frac{1}{1-\alpha}}$. $\square$

2.6 A countable set of positive Minkowski dimension

Lemma 2.6.1. *If* $E = \{0\} \cup \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, ...\}$, *then*

$$\dim_M E = \frac{1}{2}.$$

Proof. Notice that for every $n \geq 1$, we have

$$\frac{1}{n-1} - \frac{1}{n} = \frac{1}{n(n-1)} > \frac{1}{n^2} \tag{2.7}$$

Also, for every $\epsilon > 0$, there is a number n such that,

$$\frac{1}{(n+1)^2} < \epsilon \leq \frac{1}{(n)^2}$$

Note that since $\epsilon \leq \frac{1}{n^2}$, from (2.7) it follows that we need at least n intervals of length ϵ to cover $E \cap [\frac{1}{n}, 1]$. Moreover, since $\epsilon > \frac{1}{(n+1)^2}$, to cover $E \cap [0, \frac{1}{n+1}]$ we only need

$$\frac{\frac{1}{n+1}}{\epsilon} < \frac{\frac{1}{n+2}}{\frac{1}{(n+1)^2}} = n + 1 \leq 2n.$$

intervals of length ϵ. Therefore,

$$N(E, \epsilon) \leq n + \frac{\frac{1}{n}}{\epsilon} \leq n + \frac{1}{n} \cdot n^2 \leq 2n.$$

for n large.

Therefore ,

$$\begin{aligned}
\overline{\dim}_M(E) &\leq \limsup_{\epsilon \to 0} \frac{\log N(E, \epsilon)}{\log \frac{1}{\epsilon}} \\
&\leq \limsup_{\epsilon \to 0} \frac{\log(2\epsilon^{\frac{-1}{2}})}{\log(\frac{1}{\epsilon})} \\
&= \limsup_{\epsilon \to 0} \frac{\log 2 + \log \epsilon^{\frac{-1}{2}}}{\log(\frac{1}{\epsilon})} \\
&= \limsup_{\epsilon \to 0} \frac{\log 2 + \frac{1}{2} \log \epsilon^{-1}}{\log(\frac{1}{\epsilon})} \\
&= \limsup_{\epsilon \to 0} \frac{\log 2}{\log(\frac{1}{\epsilon})} + \frac{1}{2} \\
&= \frac{1}{2}.
\end{aligned}$$

Conversely,

$$N(E, \epsilon) \geq N\left(E \cap \left[0, \frac{1}{(n+1)}\right]\right) \geq \frac{\frac{1}{n+1}}{\epsilon} \geq \frac{\frac{1}{2n}}{\epsilon} \geq \frac{1}{2} \cdot \frac{1}{\sqrt{\epsilon}}.$$

So, just like above, we get $\underline{\dim}_M(E) \geq \frac{1}{2}$. Therefore, the Minkowski dimension exists and is equal to $\frac{1}{2}$. $\square$

2.7 Sets defined by digit restrictions

In this section we define a large class of Cantor sets which in particular give an example of a set whose in upper and lower Minkowski dimension are not equal.

Recall that for every point $x \in [0,1]$ its *dyadic (or binary) expansion* is the following sum

$$x = \sum_{k=1}^{\infty} \frac{x_k}{2^k}.$$

Let $S \subset \mathbb{N}$, and define

$$A_S = \left\{ x = \sum_{k \in S}^{\infty} \frac{x_k}{2^k} \mid x_k \in \{0,1\} \right\}. \tag{2.8}$$

Thus, A_S is the collection of points in $[0,1]$ with dyadic expansion such that if $k \in S$ then x_k can be either 0 or 1, while if $k \notin S$ then $x_k = 0$, or that the corresponding fraction does not appear in the sum.

Accordingly, we can construct A_S geometrically by removing certain dyadic intervals from $[0,1]$. More precisely we may proceed by performing the following steps.

Step 0.
- Let $A_{S,0} = I_0 = [0,1]$.

Step 1.
- Divide I_0 into two equal length subintervals $[0,1/2]$ and $[1/2,1]$.
- If $1 \in S$ then keep both intervals. If $1 \notin S$ keep only the left interval, $[0,1/2]$.
- Let $A_{S,1}$ be the union of the remaining intervals.

Step 2.
- Divide every interval I of length $1/2$ left after Step 1 into two equal length closed subintervals I_R and I_L.
- If $2 \in S$ then keep both intervals. If $2 \notin S$ then keep only the left inferval I_L.

- Let $A_{S,2}$ be the union of the remaining intervals.

...

<u>*Step n.*</u>

- Divide every interval I of length $1/2^{n-1}$ left after Step $n-1$ into two equal closed length subintervals I_R and I_L.
- If $n \in S$ then keep both intervals. If $n \notin S$ then keep only the leftmost I_L.
- Let $A_{S,n}$ be the union of the remaining intervals.

...

Finally, let

$$A_S = \bigcap_{i=0}^{\infty} A_{S,n}.$$

Note, that A_S is a compact set, since it is an intersection of closed subsets of $[0,1]$. It is easy to see that the two definitions of the set A_S given above are equivalent, however for computing the Minkowski dimensions of the set it is more convenient to work with the geometric definition given above. This definition allows us to show that the Minkowski dimension of A_S can be calculated using the densities of the set $S \subset \mathbb{N}$.

Definition 2.7.1. *Given $S \subset \mathbb{N}$, we define the* ***upper and lower densities of*** *S, as follows:*

$$\overline{d}(S) = \overline{\lim}_{N\to\infty} \frac{\#(S \cap \{1,...,N\})}{N}.$$
$$\underline{d}(S) = \underline{\lim}_{N\to\infty} \frac{\#(S \cap \{1,...,N\})}{N}.$$

If $\overline{d}(S) = \underline{d}(S)$, then the limit exists and is simply called the density of S and we write it as $d(S)$.

Lemma 2.7.2. *Let $S \subset \mathbb{N}$, and A_S be defined as above. Then we have*

$$\overline{\dim}_M(A_S) = \overline{d}(S). \tag{2.9}$$

$$\underline{\dim}_M(A_S) = \underline{d}(S). \tag{2.10}$$

Proof. $A_{S,n}$ is a union of $2^{s_1+s_2+...+s_n}(= 2^{\#\{S\cap\{1,...,n\}\}})$ dyadic intervals of generation n, i.e. of length $\frac{1}{2^n}$, where

$$s_i = \begin{cases} 0 & \text{if } i \notin S \\ 1 & \text{if } i \in S \end{cases}$$

Indeed, the number of dyadic intervals in $A_{S,n}$ is obtained from the corresponding number for $A_{s,n-1}$ by multiplying by 1 if $n \notin S$, or by 2 if $n \in S$. That is to write

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

where, $\chi_S(k)$ 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,

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

Therefore,

$$\log N(S, 2^{-n}) = \left(\sum_{k=1}^{n} \chi_S(k)\right) \log 2$$

Thus,

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

Similarly, $\underline{\dim}_M(A_S) = \underline{d}(S)$, thus completing the proof. □

It follows from the example above that there are Cantor sets which have different upper and lower Minkowski dimensions. For this one merely needs to choose S so that the lower and upper densities of S are not the same. In particular, Minkowski dimension does not exist. Next we show that for subset of R in fact the upper dimension can be as large as possible, while the lower dimension can be as small as we wish. As the lemma above shows, the idea would be to construct a set S such that $0 = \underline{d}(S) < \overline{d}(S) = 1$ and consider the corresponding set A_S.

Lemma 2.7.3. *There exists a set $S \subset N$ such that*

$$\underline{d_S} = 0 \text{ and } \overline{d_S} = 1$$

Proof. Consider a subset of integer that has long gaps followed by even by longer intervals, so let

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

So S consists of intervals of integers of length $((2k+1)! - (2k)!)$ followed by a gap of length $(2k+2)! - (2k+1)!$. We first estimate the lower density, as follows

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

Therefore, $\underline{d_S} = 0$. Similarly,

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

And therefore $\overline{d_S} = 1$. □

Combining the last two Lemmas (2.7.2) and (2.7.3). We immediately get the following corollary.

Corollary 2.7.4. *There is a compact set $E \subset \mathbb{R}$ such that $\overline{\dim}(E) = 1$ and $\underline{\dim}(E) = 0$.*

2.7.1 Disadvantages of Minkowski Dimension

As we saw above Minkowski dimension of a set does not necessarily exist. Another disadvantage of $\dim_M E$ is that the equality $\dim(\cup_n E_n) = \sup \dim E_n$, which one would expect to hold for a natural notion of dimension does not hold for Minkowski dimension. Indeed, as we saw

$$\frac{1}{2} = \dim_M \left(\{0\} \cup \{1, \frac{1}{2}, \frac{1}{3}, ...\} \right) \neq \sup_n \dim_M \left(\left\{ \frac{1}{n} \right\} \right) = 0.$$

Hausdorff dimension, which we will study below, behaves much nicer, in this respect. In particular, it will be shown that $\dim_H(E) \leq \underline{\dim}_M(E)$ for every set $E \subset R^N$. Therefore, in the example of corollary 2.7.4, even though $\dim_M(E)$ does not exist, Hausdorff dimension is in fact equal to 0.

Chapter 3

Hausdorff Dimension

Our next notion of dimension is Hausdorff dimension. Hausdorff dimension is probably the most useful and most commonly used notion of fractal dimension. Recall, that Minkowski dimension can only be defined for totally bounded sets, which in R^n implies that the set has to be bounded, see Definition 2.1.1. On the other hand, Hausdorff dimension has the convenience of being defined for any set. Moreover, the definition of Hausdorff dimension is based on measures, which eliminates several of the disadvantages Minkowski dimension has. Namely,

- Hausdorff dimension of every metric space exists.
- $\dim_H(E) \neq \dim_H \overline{E}$, i.e. the Hausdorff dimension of a set is not the same as the Hausdorff dimension of its closure, in general.
- $\dim_H(\cup_n E_n) = \sup_n \dim_H E_n$.

Recall, from Chapter 2, that none of properties above hold for Minkowski dimension. In particular, Corollary 2.7.4 shows that there are Cantor sets whose Minkowski dimension does not exist. The last property above implies, in particular, that Hausdorff dimension of any countable set is zero, while Example 2.6.1 shows this is also not true for Minkowski dimension.

Nevertheless, one drawback of Hausdorff dimension is that it often tends to be quite difficult to calculate or to estimate. In this chapter, after defining the Hausdorff dimension we will prove a simple yet powerful method to obtain lower bounds called the Mass Distribution principle.

3.1 Definitions

Definition 3.1.1. *Given any set E in a metric space X and $\alpha \geq 0$, we define the α -* ***dimensional Hausdorff content*** *of E, by*

$$H^{\alpha}_{\infty}(E) = \inf\left\{\sum_{i=1}^{\infty}(\mathrm{diam}(E_i))^{\alpha} | E \subset \bigcup_{i=1}^{\infty} E_i\right\} \tag{3.1}$$

where $\{E_i\}$ is a countable cover of E, and as usual $\mathrm{diam}(E_i)$ *is the diameter of E_i.*

Definition 3.1.2. ***The Hausdorff dimension*** *of a set E is defined to be*

$$\dim_H(E) = \inf\{\alpha : H^{\alpha}_{\infty}(E) = 0\}. \tag{3.2}$$

More generally, we define for every $\epsilon > 0$

$$H^{\alpha}_{\epsilon}(E) = \inf\left\{\sum_{i=1}^{\infty}(\mathrm{diam}(E_i))^{\alpha} | E \subset \bigcup_{i=1}^{\infty} E_i, \mathrm{diam}(E_i) < \epsilon\right\} \tag{3.3}$$

It turns out that $H^{\alpha}_{\epsilon}(E)$ is an outer measure.

Definition 3.1.3. *The α-* ***dimensional Hausdorff measure*** *of a set E is defined as follows,*

$$H^{\alpha}(E) = \lim_{\epsilon \to 0} H^{\alpha}_{\epsilon}(E), \tag{3.4}$$

where the limit exits because $H^{\alpha}_{\epsilon}(E)$ is a decreasing function in ϵ.

It turns out that H^α is a metric outer measure and all the Borel sets are H^α measurable for every $\alpha \geq 0$. Before continuing we recall these definitions.

Definition 3.1.4. *Let X be a nonempty set. An* ***outer measure or exterior measure*** *on X is a function $\mu^* : P(X) \to [0, \infty]$ that satisfies the following properties:*
(a) $\mu^*(\emptyset) = 0$.
(b) Monotonicity: If $E_1 \subset E_2$, then $\mu^*(E_1) \leq \mu^*(E_2)$.
(c) Countable subadditivity: If $E_1, E_2, ... \subset X$, then

$$\mu^*(\bigcup_k E_k) \leq \sum_k \mu^*(E_k).$$

Definition 3.1.5. *Let μ^* be an outer measure on a set X. Then a set $E \subseteq X$ is* ***μ^*-measurable****, or simply measurable for short, if $\forall A \subseteq X$,*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

Definition 3.1.6. *Let (X, d) be a metric space. An outer measure μ on X is called a* ***metric outer measure*** *if*

$$dist(A, B) > 0 \Longrightarrow \mu(A \cap B) = \mu(A) + \mu(B)$$

where A and B are two subset of X.

The following result is well known and can be found in[?] .

Theorem 3.1.7. *Let μ be a metric outer measure. Then all Borel sets are μ-measurable.*

The next result allows one to define the Hausdorff dimension in terms of a “critical exponent” for the Hausdorff measures, rather that for the Hausdorff content of a set.

Lemma 3.1.8. *Consider $E \subset X$. If $0 \leq \alpha < \beta < \infty$ then*

1. *if $H^\alpha(E) < \infty$ then $H^\beta(E) = 0$,*
2. *if $H^\beta(E) > 0$ then $H^\alpha(E) = \infty$.*

Proof. Let $\epsilon > 0$ be such that $H^\alpha_\epsilon(E) \leq H^\alpha(E) + 1$. Let $E \subset \cup_i E_i$ with $\dim(E_i) \leq \epsilon$ and $\sum_i \dim(E_i)^\alpha \leq H^\alpha_\epsilon(E) + 1$. Then

$$H^\beta_\epsilon(E) \leq \sum_i \operatorname{diam}(E_i)^\beta \leq \epsilon^{\beta-\alpha} \sum_i \operatorname{diam}(E_i)^\alpha \tag{3.5}$$

$$\leq \epsilon^{\beta-\alpha}[H^\alpha_\epsilon(E) + 1]. \tag{3.6}$$

Since $\beta - \alpha > 0$ we have that $\epsilon^{\beta-\alpha} \to 0$ as ϵ is decreasing. Therefore, since $H^\alpha(E) < \infty$ we have $H^\beta_\epsilon(E) \to 0$ as $\epsilon \to 0$. Part (2) is a restatement of (1) so the proof would be the same. □

Thus if we think of $H^\alpha(E)$ as a function of α, the graph of $H^\alpha(E)$ versus α shows that there is a critical value of α where $H^\alpha(E)$ jumps from ∞ to 0. This critical value is equal to the Hausdorff dimension of the set. More generally, we have, see[?] Proposition 1.2.6, the following.

Remark 3.1.9. *For every metric space E we have*

$$H^\alpha(E) = 0 \iff H^\alpha_\infty(E) = 0 \tag{3.7}$$

Therefore, Hausdorff dimension can also be defined as follows

$$\dim_H(E) = \inf\{\alpha : H^\alpha(E) = 0\} = \inf\{\alpha : H^\alpha(E) < \infty\},$$

$$= \sup\{\alpha : H^\alpha(E) > 0\} = \sup\{\alpha : H^\alpha(E) = \infty\}.$$

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 R^N, then the following properties hold, see*?

- *1.* (Monotonicity) If $E \subset F$ then $\dim_H(E) \leq \dim_H(F)$.
- 2. (Countable Stability) If $F_1, F_2, ...$ is a countable collection of sets, then

$$\dim_H \cup_{i=1}^{\infty} F_i = \sup_{1 \leq i < \infty} \{\dim_H(F_i)\}. \tag{3.8}$$

- 3. For all $E \subset R^N$,

$$\dim_H(E) \leq \underline{\dim}_M(E) \leq \overline{\dim}_M(E). \tag{3.9}$$

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

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

3. Indeed, if $B_i = B(x_i, \epsilon/2)$ are $N(E, \epsilon)$ balls of radius $\epsilon/2$ and centers x_i in E that cover E, then consider the sum

$$S_\epsilon = \sum_{i=1}^{N(E,\epsilon)} |B_i|^\alpha = N(E, \epsilon)\epsilon^\alpha = \epsilon^{\alpha - R_\epsilon},$$

where $R_\epsilon = \frac{\log N(E,\epsilon)}{\log(1/\epsilon)}$. If $\alpha > \liminf_{\epsilon \to 0} R_\epsilon = \underline{\dim}_M E$ then $\inf_{\epsilon > 0} S_\epsilon = 0$. Strict inequality in (3.9)are possible. □

Property (3.9) is very useful because it often 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.

3.3 Examples

3.3.1 Countable and open sets

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

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

Lemma 3.3.2. *If $E \subset R^N$ is open, then* $\dim_H(E) = n$.

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

3.3.2 Middle thirds Cantor set: the upper bound

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

Lemma 3.3.3. *If C is the middle thirds Cantor set then*

$$\dim_H(C) \leq \log_3 2. \tag{3.10}$$

Proof. Let $\alpha = \log_3 2$. First we will show that $\dim_H(C) \leq \alpha$. We should show that if $\beta > \alpha$, then $H^\beta(C) = 0$. Pick $n \geq 0$ and let $I_{0,1}, ..., I_{n,2^n}$ be the 2^n intervals that comprise C_n, each of length $1/3^n$ in the construction of the Cantor set from Section 2.4. Since $C \subset C_n$, this is a cover of C. We compute the β-length of the cover. It follows that

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

Since $\beta > \log_3 2$, we have $(\frac{2}{3^\beta}) < 1$ and

$$\left(\frac{2}{3^\beta}\right)^n \underset{n\to\infty}{\longrightarrow} 0.$$

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

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

3.4 Lower estimates of Hausdorff dimension

3.4.1 Mass Distribution Principle

Theorem 3.4.1 (Mass Distribution Principle)**.** *Suppose $E \subset R^N$ and $\alpha \geq 0$. If there is a non-trivial measure μ on R^n, such that $\mu(E) > 0$, and a constant $0 < A < \infty$ such that*

$$\mu\left(B(x,r)\right) \leq Ar^\alpha, \tag{3.11}$$

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

$$\dim_H(E) \geq \alpha. \tag{3.12}$$

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

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

We deduce that $\mu(U_i) \leq B\text{diam}(U_i)^\alpha$ where $B = A \cdot 2^\alpha$, that is,

$$\sum_i \text{diam}(U_i)^\alpha \geq \sum_i \frac{\mu(U_i)}{B} \geq \frac{\mu(E)}{B},$$

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

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

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

3.4.2 Middle-thirds Cantor set: lower bound

Lemma 3.4.2. *Let C be the middle thirds Cantor set. Then*

$$\dim_H(C) \geq \log_3 2. \tag{3.13}$$

Proof. We will show that there exists a measure μ on C such that there is a constant $A < \infty$ such that for every $I \subset R$ we have $\mu(I) \leq A \text{ diam}(I)^{\log_3 2}$. We will proceed in two steps. First, we will show that $\mu(I_{n,k}) \leq A \text{ diam}(I_{n,k})^{\log_3 2}$ for every n and k. Then we will generalize this and show that the same inequality holds true for any I, possibly with a different constant A.

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 $\mu(I_{n,k}) = \frac{1}{2^n}$, for every $n > 1$ and every $k \in \{1, ..., 2^n\}$. Recall from our construction in Section 2.4 that we have $\text{diam}(I_{n,k}) = \frac{1}{3^n}$.

Therefore, we have

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

We can take $A = 1$ and we obtain

$$\mu(I_{n,k}) = \mathrm{diam}(I_{n,k})^{\log_3 2}. \tag{3.14}$$

To show that in general, $\mu(I) \leq A\,\mathrm{diam}(I)^{\log_3 2}$, we will choose n so that $\frac{1}{3^n} \leq \mathrm{diam}(I) < \frac{1}{3^{n-1}}$. Since $\mathrm{diam}(I) < \frac{1}{3^{n-1}}$, and gaps between the intervals of generation $n-1$ are of length $1/3^{n-1}$, I intersects at most one $I_{n-1,k}$. Therefore, I can intersect at most two intervals of the form $I_{n,j}$ and $I_{n,j+1}$. Hence, we have $\mu(I) \leq 2\mu(I_{n,j})$ and using (3.14), we get

$$\mu(I) \leq \mu(I_{n,j}) + \mu(I_{n,j+1}) = 2 \cdot \mathrm{diam}(I_{n,j})^{\log_3 2} \leq 2 \cdot \mathrm{diam}(I)^{\log_3 2}.$$

Since this estimate holds for every $I \subset R$, the Mass Distribution Principal implies $\dim_H = \log_3 2$. □

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

$$\dim_H(C) = \log_3 2.$$