AMERICAN UNIVERSITY OF BEIRUT

ON THE MAGNITUDE OF METRIC SPACES

by
OLA GHASSAN EL KHATIB

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science
to the Department of Mathematics
of the Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
April 2015

AMERICAN UNIVERSITY OF BEIRUT

ON THE MAGNITUDE OF METRIC SPACES

by

Ola Ghassan El Khatib

Approved by:

Farouk Abi-Khuzam

Professor Farouk Abi-Khuzam, Professor Mathematics — Advisor

Professor Florian Jacques Bertrand, Assistant Professor Mathematics — Member of Committee

Professor Tamer Tlas, Assistant Professor Mathematics — Member of Committee

Date of thesis defense: April 27, 2015

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: El Khatib Ola Ghassan
Last First Middle

☑ Master's Thesis ○ Master's Project ○ Doctoral Dissertation

☐ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☑ I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project,** to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

OKhatib 30.04.2015

Signature Date

Acknowledgements

Foremost, I would like to express my deepest gratitude to my advisor Prof. Faruk Abi-Khuzam for his continuous support, insightful comments and motivation throughout my thesis work. From him, I have learned many precious values not only in academic work, but in life in general.

Sincere thanks to my family (especially my sister) and friends for their encouragement in every step in my Masters studies. Also, special thanks to my friend Nagham El Chaar who made the graduate studies experience a memorable one.

Above all, I dedicate my profound gratefulness to GOD for His endless blessings.

AN ABSTRACT OF THE THESIS OF

Ola Ghassan El Khatib for Master of Science
Major: Mathematics

Title: On the magnitude of metric spaces

In this thesis, we consider the magnitude of finite metric spaces and that of compact subsets of Euclidean spaces. We will consider an extension of the definition of finite spaces to compact metric spaces using three different approaches. Our contribution to this work consists in giving new proofs for the magnitude of segments and general compact sets of $\mathbb{R}$. In addition, we give a presentation of the proof that the magnitude dimension of compact subsets of Euclidean space is equal to the Minkowski dimension.

Contents

Chapter 1

Introduction

The concept of magnitude of a metric space was introduced by Tom Leinster in 2008. The idea comes from analogies with category theory and the general notion of size such as cardinality of sets, dimension of vector spaces, Euler characteristic of topological spaces, and entropy of probability spaces.

There is a an analogy that can be drawn between metric spaces and categories; as categories have objects (a, b, c...), metric spaces have points (a, b, c...); as any two objects in categories are related by set of maps between them (Hom(a,b)), any two points in a metric metric space A are related by a real non-negative number representing the distance between them $\Big($d(a,b) such that $d : A \times A \to [0, \infty]\Big)$; as there is the notion of composition in category: $Hom(a, b) \times Hom(b, c) \to Hom(b, c)$, there is the triangle inequality in metric spaces that relates distances: $d(a, b) + d(b, c) \geq d(a, c)$. This analogy lends some justification to the introduction of the concept of magnitude, but it will not be perused in the present thesis. Our emphasis will be on the problem of computation of the magnitude of certain metric spaces. Despite its apparent simplicity this problem continues to be open for a large number of metric spaces.

Consider a metric space with n-points which are a distance d apart. When d is very small, the magnitude is just greater than one i.e there's effectively only one point. The magnitude will increase with the distance d such that when d is large enough, the magnitude will be just less than n and therefore there are effectively n points in the space. The magnitude of a finite metric space was introduced as the 'effective number of species' in the biodiversity literature.

One of the main motivations that are proposed to justify the study of magnitude are the conjectures made by Leinster and Meckes to the effect that magnitude captures various aspects of the geometry of the metric space. As in most conjectures, one needs to study specific examples and compute if possible the magnitude in some non-trivial cases. It should be noted that the number of examples in the literature is rather limited and it remains an important problem to calculate the magnitude of many natural examples as mentioned in [3].

Our contribution in this thesis lies in finding the magnitude of compact subsets of $\mathbb{R}$, where we show that for all K non-empty compact subsets of $\mathbb{R}$, the magnitude is given by

$$|K| = 1 + \frac{l(K)}{2} + \sum_{I} tanh \frac{l(I)}{2}$$

where the sum extends over the bounded open intervals of K^c and $l(E)$ is the Lebegue measure of the set E.

The second chapter of this thesis will introduce the magnitude of finite metric spaces with some methods of calculation. It will be noticed that not all spaces admit a well-defined magnitude and this leads to the introduction of homogeneous spaces and positive definite spaces. In chapter three, the magnitude definition will be extended to compact metric spaces using three different approaches. Finally in chapter four, the proof by Meckes of the relationship between diversity, Minkowski and magnitude dimensions will be presented. There, it is shown that the diversity and Minkowski dimensions are the same for compact metric spaces and that Minkowski and magnitude dimensions are equal for Euclidean subspaces.

Chapter 2

Finite Metric Spaces

2.1 Definition of Magnitude

Given a finite metric space A, we define the following based on [3]

- **Similarity matrix** of exponentiated distances $\zeta_A \in \mathbb{R}^{A \times A}$ is defined by

$$\zeta_A(a, b) = e^{-d(a,b)}(a, b \in A).$$

  In case ζ_A is invertible, we say that A has a **Mobius inversion** and the **Mobius matrix** $\mu_A = \zeta_A^{-1}$.

- **Weighting** on A is a function $w : A \to \mathbb{R}$ such that $\sum_b \zeta_A(a, b)w(b) = 1$ for all a $\in$ A.

- The space A has **magnitude** if there exists at least one weighting. The magnitude is defined by

$$|A| = \sum_a w(a)$$

  for any weighting w.

There are several things to note from this definition:

- The magnitude is not defined for all spaces since the weighting may not exist. For L_p spaces where p> 2 there exists finite subsets where the magnitude is not defined as mentioned in [3].

- The weighting is not necessarily positive.

- The weighting may not be unique.

- The magnitude is **unique** and is **independent of the choice of the weighting**.

  Suppose w and w' are two weightings then,

$$\sum_a w_a = \sum_a w_a \sum_b e^{-d(a,b)} {w'}_b = \sum_b \sum_a w_a e^{-d(a,b)} {w'}_b = \sum_b {w'}_b$$

  this is true since $d(a, b) = d(b, a)$.

- If ζ_A is invertible, then the weighting is unique. The weighting $w(a) = \sum_b \mu(a,b)$ and $|A| = \sum_{a,b} \mu_A(a,b)$

Example 2.1. *i. The magnitude of an empty set is 0.*

ii. The magnitude of a space with only one point is 1.

iii. Let A be a space of two points a and b distance d apart. Then, similarity matrix $\zeta_A = \begin{bmatrix} 1 & e^{-d} \\ e^{-d} & 1 \end{bmatrix}$.

Then det $(\zeta_A) = 1 - e^{-2d} \neq 0$ and therefore ζ_A is invertible. So that

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

$$\begin{aligned} |A| &= \sum_a w(a) \\ &= \frac{2-2e^{-d}}{1-e^{-2d}} = \frac{2}{1+e^{-d}} \\ &= \frac{e^{d/2}+e^{-d/2}}{e^{d/2}+e^{-d/2}} + \frac{e^{d/2}-e^{-d/2}}{e^{d/2}+e^{-d/2}} \\ &= 1 + tanh(d/2). \end{aligned}$$

Notice that when d is very small $|A|$ approaches 1 which is the magnitude of the space with one point. Furthermore, as d gets larger the points get more separated and $\lim_{d\to\infty} |A| = 2$ which is the cardinality of the space A.

iv. A is a space of three points x_1, x_2 and x_3 such that $d(x_1,x_2) = d_1$, $d(x_2,x_3) = d_2$ and $d(x_1,x_3) = d_3$.

The similarity matrix $\zeta_A = \begin{bmatrix} 1 & e^{-d_1} & e^{-d_3} \\ e^{-d_1} & 1 & e^{-d_2} \\ e^{-d_3} & e^{-d_2} & 1 \end{bmatrix}$

$$\begin{aligned} det(\zeta_A) &= 1 - e^{-2d_1} - e^{-2d_2} - e^{-2d_3} + 2e^{-(d_1+d_2+d_3)} \\ &= (1-e^{-d_1})(1-e^{-d_2})(1-e^{-d_3}) + (1-e^{-d_1})(e^{-d_1} - e^{-d_2-d_3}) + (1-e^{-d_2})(e^{-d_2} - e^{-d_1-d_3}) \\ &\quad + (1-e^{-d_3})(e^{-d_3} - e^{-d_1-d_2}) \end{aligned}$$

$(1-e^{-d}) > 0$ *and* $d_2 + d_3 \geq d_1$ *by the triangle inequality therefore* $(e^{-d_1} - e^{-d_2-d_3}) \geq 0$ *and* $det(\zeta_A) > 0$ *and so* ζ_A *is invertible.*

$$\mu_A = \zeta_A^{-1}(a,b) = det(\zeta_A)\times \begin{bmatrix} 1-e^{-d_2} & -(e^{-d_1}-e^{-(d_2+d_3)}) & -(e^{-(d_1+d_2)}-e^{-d_3}) \\ -(e^{-d_1}-e^{-(d_2+d_3)}) & 1-e^{-d_3} & -(e^{-(d_1+d_3)}-e^{-d_2}) \\ -(e^{-(d_1+d_2)}-e^{-d_3}) & -(e^{-(d_1+d_3)}-e^{-d_2}) & 1-e^{-d_1} \end{bmatrix}$$

$$|A| = \Sigma_a w(a) = \frac{3-e^{-2d_1}-e^{-2d_2}-e^{-2d_3}-2(e^{-d_1}+e^{-d_2}+e^{-d_3}-e^{-(d_1+d_2)}-e^{-(d_1+d_3)}-e^{-(d_2+d_3)})}{1-e^{-2d_1}-e^{-2d_2}-e^{-2d_3}+2e^{-(d_1+d_2+d_3)}}$$

If $d_1 = d_2 = d_3 = d$, *then* $|A| = \dfrac{3-6e^{-d}+3e^{-2d}}{1+2e^{-3d}-3e^{-2d}}$.

v. *Let A be a metric space such that $d(a,b)=\infty$ for all $a \neq b$ in A. Then the matrix ζ_A is the identity matrix, each point has a weighting of 1 and $|A| = \#A$ [3].*

Since not all spaces admit a magnitude, we need to study and categorize the types of spaces with a well-defined magnitude. In example 2.1 (iv) it was shown that all spaces with 3 points have magnitude, and Meckes [26, Theorem 3.6] proved that spaces with four points have magnitude. But spaces with cardinality 5 or more may not admit a defined magnitude.

2.2 Magnitude of a Union of Metric Spaces

When considering the magnitude of a union of metric spaces, we cannot handle it as the cardinality of sets is handled where we have the inclusion-exclusion principle. So $|A \cup B|$ is not generally deduced from $|A|, |B|$ and $|A \cap B|$. However, for unions of special types, we have the inclusion-exclusion formula for magnitude. We will consider the magnitudes of such unions as discussed by Leinster in [3].

Definition 2.2. *Let X be a metric space and $A, B \subset X$. Then* ***A projects to B*** *if for all $a \in A$ there exists $\pi(a) \in A \cap B$ such that for all $b \in B$,*

$$d(a,b) = d(a,\pi(a)) + d(\pi(a),b).$$

Clearly then in such a case $d(a,\pi(a)) = inf_{b\in B} d(a,b)$.

Proposition 2.3. *Let X be a metric space and $A, B \subset X$. Suppose that A projects to B and B projects to A. If A and B have magnitude then so does $A \cup B$ with*

$$|A \cup B| = |A| + |B| - |A \cap B|$$

If w_A, w_B and $w_{A\cap B}$ are weightings for A, B and $A \cap B$ respectively, then the weighting $w_{A\cup B}$ for $A \cup B$ is defined by

$$w_{A\cup B}(x) = \begin{cases} w_A(x) & \text{if } x \in A\backslash B \\ w_B(x) & \text{if } x \in B\backslash A \\ w_A(x) + w_B(x) + w_{A\cap B}(x) & \text{if } x \in A \cap B. \end{cases}$$

Proof. First check if the suggested weighting satisfies the weighting condition. Consider a in $A\backslash B$, since A projects to B, then for all a $\in$ A theres exists $\pi(a) \in A \cap B$ such that for all b $\in$ B $d(a,b) = d(a,\pi(a)) + d(\pi(a),b)$.

$$\begin{aligned}
\sum_{x\in A\cup B} \zeta(a,x)w(x) &= \sum_{a'\in A} \zeta(a,a')w_A(a') + \sum_{b\in B} \zeta(a,b)w_B(b) - \sum_{c\in A\cap B} \zeta(a,c)w_{A\cap B}(c) \\
&= \sum_{a'\in A} e^{-d(a,a')}w_A(a') + \sum_{b\in B} e^{-d(a,b)}w_B(b) - \sum_{c\in A\cap B} e^{-d(a,c)}w_{A\cap B}(c) \\
&= 1 + \sum_{b\in B} e^{-d(a,\pi(a))}e^{-d(\pi(a),b)}w_B(b) - \sum_{c\in A\cap B} e^{-d(a,\pi(a))}e^{-d(\pi(a),c)}w_{A\cap B}(c) \\
&= 1 + e^{-d(a,\pi(a))}(1-1) \\
&= 1
\end{aligned}$$

Similarly if a $\in B\backslash A$ and $A \cap B$.

Now since we have the weighting, we can get the magnitude.

$$\begin{aligned}
|A \cup B| &= \sum_{a\in A\cup B} w_{A\cup B}(x) \\
&= \sum_{a\in A} w_A(a) + \sum_{b\in B} w_B(b) + \sum_{c\in A\cap B} w_{A\cap B}(c) \\
&= |A| + |B| - |A \cap B|
\end{aligned}$$

□

Corollary 2.4. *Let X be a metric space and $A, B \subset X$. Suppose that A intersects B in only one point c such that for all a$\in$ A and $b \in B$,*

$$d(a,b) = d(a,c) + d(c,b)$$

and A and B both have magnitudes. Then $|A \cup B| = |A| + |B| - 1$.

Proof. From proposition 2.3, $|A\cup B| = |A|+|B|-|A\cap B|$ and since $A\cap B = \{c\}$ and the magnitude of a single point is 1. Then $|A \cup B| = |A| + |B| - 1$. □

Corollary 2.5. *Every finite subspace of $\mathbb{R}$ has a Mobius inversion. Let A=$\{a_0, ..., a_n\}$ where $n \geq 0$, and*

$$0 \leq a_0 < a_1 < ... < a_n$$

be a finite subset of $\mathbb{R}$. Then the magnitude of A is given by

$$|A| = 1 + \sum_{i=1}^{n} tanh(\frac{d_i}{2}), d_i = a_i - a_{i-1},$$

where the sum is taken to be empty if n=0.

Proof. For n=1, the magnitude of $A = \{a_0, a_1\}$ was calculated in example 2.1 (iii) to be

$$|A| = 1 + tanh(d/2) = 1 + \sum_{i=1}^{1} tanh(\frac{d_i}{2}).$$

For n=2, suppose $A_1 = \{a_0, a_1\}$, $A_2 = \{a_1, a_2\}$, $d_1 = a_1 - a_0$ and $d_2 = a_2 - a_1$ with $a_0 < a_1 < a_2$ and $A = A_1 \cup A_2$.

Now, $A_1 \cap A_2 = a_1$, then by Collorary 2.4,

$$\begin{aligned} |A| &= |A_1| + |A_2| - 1 \\ &= 1 + tanh(d_1/2) + 1 + tanh(d_2/2) - 1 \\ &= 1 + \sum_{i=1}^{2} tanh(\frac{d_i}{2}). \end{aligned}$$

Suppose the formula for the magnitude is true for $\#A = n$.

Let $A = A_n \cup A_{n+1}$ where $A_n = \{a_0, a_1, ..., a_n\}$, $A_{n+1} = \{a_n, a_{n+1}\}$ and $d_i = a_i - a_{i-1}$.

Then

$$\begin{aligned} |A| &= |A_n| + |A_{n+1}| - 1 \\ &= 1 + \sum_{i=1}^{n} tanh(\frac{d_i}{2}) + 1 + tanh(d_{i+1}/2) - 1 \\ &= 1 + \sum_{i=1}^{n+1} tanh(\frac{d_i}{2}). \end{aligned}$$

Hence $|A| = 1 + \sum_{i=1}^{n} tanh(\frac{d_i}{2})$ where $d_i = a_i - a_{i-1}$. □

2.3 Homogeneous Metric Spaces

Definition 2.6. *A metric space is said to be* ***homogeneous*** *if its isometry group acts transitively on points.*

Proposition 2.7. *The Speyer Formula*

Every homogeneous finite metric space A possesses a weighting where all the points have the same weight given by

$$w_a = \frac{1}{\sum_b e^{-d(a,b)}}$$

for any $b \in A$. Thus the magnitude is well defined and given by:

$$|A| = \frac{\#A}{\sum_b e^{-d(a,b)}}$$

for any $b \in A$.

Proof. Since the space possesses a transitive group action, then $\sum_b e^{-d(a,b)}$ is the same for every $a \in A$. Therefore, $w_a = \dfrac{1}{\sum_b e^{-d(a,b)}}$ is the same for all $a \in A$.

Now we have to check if this w_a satisfies the weighting condition;

$$\sum_a \Big(\frac{1}{\sum_c e^{-d(c,a)}}\Big) e^{-d(a,b)} = \frac{\sum_a e^{-d(a,b)}}{\sum_c e^{-d(c,a)}} = 1,$$

$$|A| = \sum_a w(a) = \frac{\#A}{\sum_b e^{-d(a,b)}}.$$

□

Example 2.8. *Leinster in [3] provided an example of homogeneous spaces. Given an undirected graph G and $t \in (0, \infty]$, there exists a metric space tG where the distances are the minimal edge lengths (a single edge has length t) and the points of the space are the vertices of the graph. Then*

$$|tK_n| = \frac{n}{1 + (n-1)e^{-t}}.$$

The magnitude of a homogeneous space is the reciprocal of the mean similarity since $e^{-d(a,b)}$ reflects the similarity and closeness of the points a and b[3].

2.4 Magnitude Function

Definition 2.9. *Let A be a metric space and $t \in (0, \infty)$. Then tA denotes the metric space with the same points as A and $d_{tA}(a, b) = td_A(a, b)$ where $a, b \in A$ [3].*

Definition 2.10. *Let A be a metric space. The **magnitude function** of A is the partially defined function $t \mapsto |tA|$, defined for all $t \in (0, \infty)$ for which tA has magnitude [3].*

Metric spaces possess many interesting and crucial features that allow us to uncover additional properties of the studied space. The most important feature of metric spaces is that they can be re-scaled, and examining the magnitude of the re-scaled space $|tA|$ will give more information about the space than studying the magnitude $|A|$ of the space itself. In addition, an interesting fact about re-scaling a metric space is that it does not behave in a predictable manner as the other invariants of the metric space.

It will be shown in the next section that a finite metric with sufficiently separated points has a well-defined magnitude.

Example 2.11. *Consider the finite metric A consisting of two points. It was shown in Example 2.1 (iii) that $|A| = 1 + tanh(d/2)$. The magnitude function of A is $t \mapsto 1 + tanh(td/2)$.*

Proposition 2.12. *Let A be a finite metric space. Then $|tA| \to \#A$ as $t \to \infty$.*

Proof. First note that $\zeta_{tA} \to$ Identity matrix as $t \to \infty$. Then ζ_{tA} is invertible for very large t and thus admits a magnitude.

$$\lim_{t\to\infty} |tA| = |\lim_{t\to\infty} \zeta_{tA}| = |IdentityMatrix| = \#A.$$

□

So for any finite metric space, when the distances between its points are taken to infinity, the magnitude approaches the cardinality of the space. However, the behavior when $t \to 0$ is unpredictable and the magnitude for small values of t may not be defined or even negative.

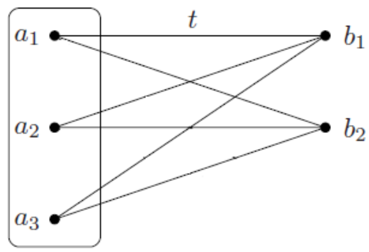

Figure 2.1: $tK_{3,2}$

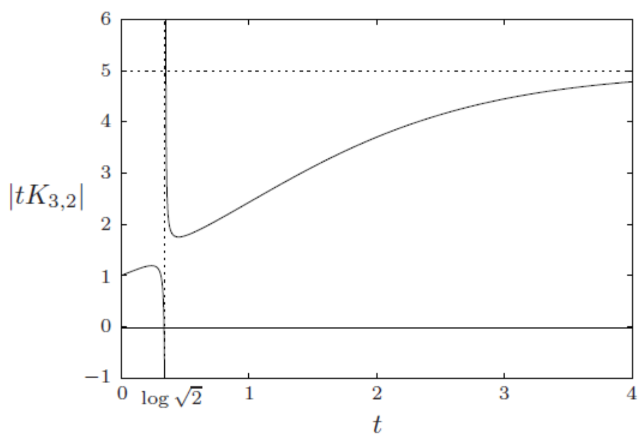

Figure 2.2: The magnitude function of $tK_{3,2}$

Example 2.13. *In [3] an example was given of a space where the magnitude behaves strangely for small values of t. The space of the graphs $K_{3,2}$ (see figure 2.1), has magnitude given by*

$$|tK_{3,2}| = \frac{5 - 7e^{-t}}{(1 + e^{-t})(1 - 2e^{-2t})}$$

Notice that for $t = log\sqrt{2}$ the magnitude is not defined as shown in figure 2.2 and $\lim_{t\to\infty} = 5$ which is the cardinality of $K_{3,2}$.

For some values of t the magnitude is negative, others it is greater than the cardinality. Notice also that $(log\sqrt{2})K_{3,2}$ is a subspace of $(log\sqrt{2})K_{3,3}$ whose magnitude is defined since it is homogeneous. Hence a space with magnitude can have a subspace whose magnitude is not defined.

2.5 Positive Definite Metric Spaces

In this section, we will show that all positive definite metric spaces (PDMS) have a well-defined magnitude. This fact is very helpful when trying to show that a certain space has magnitude, it is enough to show that it is positive definite.

Definition 2.14. *A finite metric space* ***positive definite*** *if the matrix ζ_A is positive definite.*

In [3], the following facts were proved:

Proposition 2.15. *i. A PDMS has a Mobius inversion, therefore it has a unique weighting.*

ii. A subspace of a PDMS is positive definite.

iii. The magnitude of a PDMS is increasing with respect to inclusion. i.e. if A is positive definite and $B \subseteq A$ then $|B| \leq |A|$.

iv. If A is a non-empty positive definite space then $|A| \geq 1$.

We describe a type of spaces which are positive definite and therefore their magnitude is defined.

Definition 2.16. *A finite metric space A is* ***scattered*** *if its points are sufficiently separated such that $d(a,b) > log(n-1)$ where $n = \#A$.*

Proposition 2.17. *Every scattered finite metric space has a well-defined magnitude.*

Proof. We will show that each scattered finite metric space is positive definite. Let A be such a space, it is enough to show that the similarity matrix ζ_A is positive definite.

So we need to show that a real $n \times n$ matrix ζ with $\zeta(i,i) = 1 \forall i$ and $0 \leq \zeta(i,j) < \frac{1}{n-1}$ for all $i \neq j$ is positive definite i.e. $x^T \zeta x \geq 0 forall x \in \mathbb{R}^n$ and equality holds if and only if x=0.

$x^T \zeta x = \sum_i x_i^2 + \sum_{i \neq j} \zeta_{ij} x_i x_j$

But $d(i,j) > ln(n-1)$ so $\zeta_{ij} = e^{-d(i,j)} > \frac{-1}{n-1}$

So $x^T \zeta x \geq \sum_i x_i^2 - \frac{1}{n-1} \sum_{i \neq j} |x_i|.|x_j|$

Note that

$$\begin{aligned}
\frac{1}{2(n-1)} \sum_{i \neq j} (|x_i| - |x_j|)^2 &= \frac{1}{2(n-1)} \sum_{i \neq j} (|x_i|^2 - 2|x_i||x_j| + |x_j|^2) \\
&= \frac{1}{2(n-1)} ((n-1) \sum_i (|x_i|^2) + (n-1) \sum_j |x_j|^2)) - \frac{1}{n-1} \sum_{i \neq j} |x_i||x_j| \\
&= \sum_i |x_i|^2 - \frac{1}{n-1} \sum_{i \neq j} |x_i||x_j|
\end{aligned}$$

So $x^T \zeta x \geq \frac{1}{2(n-1)} \sum_{i \neq j} |x_i||x_j| \geq 0$

If $x^T \zeta x = 0$ then all the inequalities above are equalities and $|x_i| = k \ \forall i$. Since $\zeta_A < \frac{1}{n-1}$ then $k = 0$ and hence $x = 0$. So ζ_A is positive definite and A is positive definite therefore it admits a well-defined magnitude. □

The above proof also shows that every finite metric space can be scaled up by a certain factor so that it becomes positive definite and its magnitude becomes well defined.

It was shown in [3] that every finite subspace of Euclidean space is positive definite and have a well defined magnitude.

Chapter 3

Compact Metric Spaces

In this chapter, we consider the question of extending the notion of magnitude from finite to infinite metric spaces. There are three different approaches to this and so three different ways to calculate the magnitude of an infinite metric space A.

The first approach is to define $|A|$ as the sup of the magnitude of the finite subspaces in A. This definition is not satisfactory for all types of metric spaces because the magnitude is not monotone with respect to inclusion for all spaces. As shown in example 2.13 a subspace can have greater magnitude than the whole space. However, this definition works for all compact PDMS.

The second approach is to consider a sequence of finite subspaces A_k which approximates A in the Hausdorff distance and then define $|A| = \lim_{k\to\infty} |A_k|$. This definition is not satisfactory as well since it is not clear ahead of time whether this limit is independent of the approximation chosen. It was proved by Meckes in [5] that for PDMS which includes all subspaces of Euclidean spaces, this method gives a correct definition of the magnitude irrespective of the approximation chosen. The advantage of this method is that it can be computerized to calculate magnitudes of spaces.

Definition 3.1. *For A and $B \subseteq \mathbb{R}^n$, the* ***Hausdorff distance*** *d(A,B)is defined by*

$$d(A,B) = max\Big(\sup_{a\in A} d(a,B), \sup_{b\in B} d(b,A)\Big).$$

Convergence in the Hausdorff metric sense: A sequence of compact subsets A_k of $\mathbb{R}^n$ converges to a set A if $d(A_k, A) \to 0$ as $k \to \infty$.

The third approach is to try to adapt the definite of magnitude to apply to infinite spaces directly by using weight measures and replacing sums with integrals. However, this method is applicable only when weight measures exist.

It is to be remarked that the three approaches yield the same results for PDMS and all compact subsets of the Euclidean spaces.

3.1 First Approach

Definition 3.2. *An infinite metric space is* ***positive definite*** *if every finite subspace is positive definite. The* ***magnitude*** *of a compact positive definite space A is*

$$|A| = sup\{|B| : B \text{ is a finite subspace of } A\} \in [0, \infty]$$

It is a direct consequence of monotonicity of magnitude of compact PDMS. Since, consider $K \subset X$ two compact PDMS and B is any finite subset of K then $|B| \leq |X|$ (Proposition 2.14) and hence $|K| \leq |X|$.

In order to calculate the magnitude of segments in $\mathbb{R}$ using the first approach, we need to explain how a sequence of points converges to the segment. In many references, this convergence is defined in terms of the Hausdorff convergence, however we are able to use another method and give simpler proof [1].

Since the magnitude of a singleton is 1, the magnitude of any non-empty compact subset of $\mathbb{R}$ is at least 1. Also, if A_k is a decreasing sequence of non-empty compact subsets of $\mathbb{R}$, then $\lim_{k\to\infty} |A_k|$ exists as a non-negative real number and, since the intersection is a compact set we have $|\cap_{k=1}^{\infty} A_k| \leq |A_k|$ for every k and so

$$|\cap_{k=1}^{\infty} A_k| \leq \lim_{k\to\infty} |A_k|$$

This raises the question as to whether equality holds in the above inequality. In order to answer this, we start with a particular case of a nested sequence of intervals $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$. Then a_n increases to a limit a, and b_n decreases to a limit b, and $\cap[a_n, b_n] = [a, b]$. Since, as we show next, the magnitude of $[a_n, b_n]$ is $1+\frac{b_n-a_n}{2}$, it will follow immediately that $|[a, b]| = 1 + \dfrac{b-a}{2} = \lim_{n\to\infty}\left(1+\dfrac{b_n-a_n}{2}\right) = \lim_{n\to\infty} |[a_n, b_n]|$ The formula for the magnitude of a compact segment just used is obtained using the following result on the magnitude of a finite subset of $\mathbb{R}$.

3.1.1 The magnitude of a segment in $\mathbb{R}$

Collorary 2.5 makes it possible to calculate the magnitude of any compact subset of $\mathbb{R}$. Indeed, if K is a non-empty compact subset of $\mathbb{R}$, then its complement K^c is the union of an at most countable number of open intervals I, two of which are unbounded. We shall show that the magnitude of K is given explicitly by a formula involving its Lebegue measure and the measure of the bounded open intervals in its complement. We start by determining the magnitude of a compact segment $L_l = [0, l]$ in $\mathbb{R}$.

Theorem 3.3. *Let L_l be the segment $[0, l] \subset \mathbb{R}$, then the magnitude of L_l is given by*

$$|L_l| = 1 + \frac{l}{2}$$

Proof. Since L_l is a compact sunset of $\mathbb{R}$, its magnitude $|L_l|$ is given by

$$|L_l| = sup\{|X| : \text{X is a finite subset } L_l\}.$$

Let X be a finite subset of L_l. Then X=$\{a_0, a_1, ..., a_n\}$ where

$$0 \leq a_0 < a_1 < ... < a_n \leq l.$$

Then

$$|X| = 1 + \sum_{i=1}^{n} tanh\frac{d_i}{2}, d_i = a_i - a_{i-1}.$$

The function tanh is concave and increasing on L_l so that

$$\sum_{i=1}^{n} tanh\frac{d_i}{2} \leq ntanh\Big(\frac{\sum_{i=1}^{n} d_i/2}{n}\Big) = ntanh\Big(\frac{a_n - a_0}{2n}\Big) \leq ntanh\frac{l}{2n} \leq \frac{l}{2}.$$

This gives $1 + \frac{l}{2}$ as an upper bound for $|X|$, and it follows that

$$|L_l| \leq 1 + \frac{l}{2}.$$

On the other hand, for an integer $n \geq 1$, put

$$a_i = \frac{il}{n}, i = 0, 1, ..., n$$

and take the particular finite subset $X = \{a_0, a_1, ..., a_n\}$ of L_l. Then

$$|L_l| \geq |X| = 1 + ntanh\frac{l}{2n},$$

and so

$$|L_l| \geq 1 + \lim_{n\to\infty} ntanh\frac{l}{2n} = 1 + \frac{l}{2}.$$

Hence, the magnitude of the segment is

$$|L_l| = 1 + \frac{l}{2}.$$

□

3.1.2 The magnitude of a a union of two disjoint closed segments

Let $K = [0, l] \cup [a, b]$ with $0 < l < a < b$. To find the magnitude of K, take first a finite subset X of K. Since the addition of points to a finite set increases its magnitude, we may assume X has been augmented so that

$$X = \{a_0, a_1, ..., a_n\} \cup \{b_0, b_1, ..., b_n\}$$

with $a_i \in [0, l], b_j \in [a, b]$, and $a_n = l$, and $b_0 = a$. Then we have

$$|X| = 1 + \sum_{i=1}^{n} tanh\frac{d_i}{2} + tanh\frac{a-l}{2} + \sum_{i=1}^{m} tanh\frac{d'_i}{2}, d_i = a_i - a_{i-1}, d'_i = b_i - b_{i-1}$$

$$\leq 1 + \frac{l}{2} + tanh\frac{a-l}{2} + \frac{b-a}{2},$$

and hence

$$|K| \leq 1 + \frac{l}{2} + \frac{b-a}{2} + tanh\frac{a-l}{2}.$$

To get an inequality in the opposite direction, take points $a_i \in [0, l]$, and $b_j \in [a, b]$ equally spaced in their respective intervals as before thereby obtaining $|K| \geq 1 + \frac{l}{2} + \frac{b-a}{2} + tanh\frac{a-l}{2}$, and hence equally holds.

This method obviously extends to a finite number of disjoint compact intervals of $\mathbb{R}$ and we obtain with the notation $l(E)$ equals the Lebegue measure of E.

Theorem 3.4. *Suppose K is a non-empty compact subset of $\mathbb{R}$, and K^c has a finite number of bounded open component intervals, then*

$$|K| = 1 + \frac{l(K)}{2} + \sum_I tanh\frac{l(I)}{2}$$

where the sum extends over the bounded intervals of K^c.

In fact, the restriction that K^c has a finite number of bounded components is not necessary and we have the more general results

Theorem 3.5. *Suppose K is a non-empty compact subset of $\mathbb{R}$, then*

$$|K| = 1 + \frac{l(K)}{2} + \sum_I tanh\frac{l(I)}{2}$$

where the sum extends over the bounded open intervals of K^c.

3.1.3 The magnitude of the cantor ternary set

Theorem 3.6. *The magnitude of a cantor ternary set K is given by*

$$|K| = 1 + \frac{1}{2}\sum_{j=1}^{\infty} 2^j tanh\frac{l}{2.3^j}.$$

Proof. This is the set K obtained by removing the open middle thirds starting with the interval $[0, l]$. So in the first step we remove one interval of length $\frac{l}{3}$ and end up with a compact set A_1 whose magnitude is

$$|A_1| = 1 + \frac{1}{2}(l - \frac{l}{3}) + tanh\frac{l}{2 \times 3}.$$

In the second step we remove two intervals each of length $\frac{l}{3^2}$ and end up with a compact set $A_2 \subset A_1$ whose magnitude is

$$|A_2| = 1 + \frac{1}{2}\{l - \frac{l}{3} - 2\frac{l}{3^2}\} + tanh\frac{l}{2 \times 3} + 2tanh\frac{l}{2 \times 3^2}.$$

Continuing we arrive at the set A_k with

$$|A_k| = 1 + \frac{1}{2}\left(l - \sum_{i=0}^{k} 2^i.\frac{l}{3^{i+1}}\right) + \sum_{i=0}^{k} 2^j tanh\frac{l}{2.3^{j+1}}.$$

In particular, since K is the intersection $\cap A_k$, we obtain

$$|K| \leq \lim_{k\to\infty} |A_k| = 1 + \sum_{j=0}^{\infty} 2^j tanh \frac{l}{2.3^{j+1}} = 1 + \frac{1}{2}\sum_{j=1}^{\infty} 2^j tanh \frac{l}{2.3^j}.$$

But in fact using theorem 3.5, equality holds. □

3.2 Second Approach

3.2.1 The magnitude of compact subspaces of $\mathbb{R}$

Proposition 3.7. *Let A be a compact subspace $\mathbb{R}$ then*

$$|A| = \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x, A)dx$$

where $d(x, A) = inf_{a\in A} d(x, a)$.

Proof. We first use induction to prove the formula for finite metric spaces then extend it to compact spaces. Let A be a finite subspace of $\mathbb{R}$ with $\#A = n$.

If A is a singleton, then $\frac{1}{2}\int_{-\infty}^{\infty} sech^2(x)dx = \frac{1}{2}(2) = 1 = |A|$.

Suppose $A = \{a, b\}$, then

$$d(x, A) = \begin{cases} a - x & \text{if } x < a \\ x - a & \text{if } a \leq x < \frac{a+b}{2} \\ b - x & \text{if } \frac{a+b}{2} \leq x \leq b \\ x - b & \text{if } x > b. \end{cases}$$

$$\begin{aligned} |A| &= \frac{1}{2}\Big[\int_{-\infty}^{a} sech^2(a-x)dx + \int_{a}^{\frac{a+b}{2}} sech^2(x-a)dx + \int_{\frac{a+b}{2}}^{b} sech^2(b-x)dx + \int_{b}^{\infty} sech^2(x-b)dx\Big] \\ &= \frac{1}{2}\Big[tanh(\frac{b-a}{2}) + tanh(\frac{b-a}{2}) + 2\Big] \\ &= 1 + tanh(\frac{b-a}{2}). \end{aligned}$$

Note that this agrees with the results of example 2.1 (iii).

In general, let $B = \{a_0, a_1, a_2 ... a_{n-1}\}$ and $C = \{a_{n-1}, a_n\}$, and A=$B \cup C$.

$\frac{1}{2}\int_{\mathbb{R}} sech^2 d(x, A)dx = \frac{1}{2}\Big[\int_{-\infty}^{a_{n-1}} sech^2 d(x, B) + \int_{a_{n-1}}^{\infty} sech^2 d(x, C)\Big]$

$\int_{\mathbb{R}} sech^2 d(x, B) = \int_{-\infty}^{a_{n-1}} sech^2 d(x, B) + \int_{a_{n-1}}^{\infty} sech^2 d(x, B)dx$

$\int_{a_{n-1}}^{\infty} sech^2 d(x, B)dx = tanh(x - a_{n-1}) = 1$

$$\int_{\mathbb{R}} sech^2 d(x,B)dx - 1 = \int_{-\infty}^{a_{n-1}} sech^2 d(x,B)dx$$

$$\begin{aligned}
\frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,A)dx &= \frac{1}{2}\Big[\int_{\mathbb{R}} sech^2 d(x,B)dx - 1 + \int_{\mathbb{R}} sech^2 d(x,C)dx - 1\Big] \\
&= \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,B)dx + \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,C)dx - 1 \\
&= |B| + |C| - 1.
\end{aligned}$$

Next, we need to consider compact subsets of $\mathbb{R}$, let A be such a subset.

We know that $|A| = sup\{|B| = \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,B)dx$: B is a finite subset of A$\}$.

Since $sech^2$ is a decreasing function on $[0,\infty)$ and $d(x,B) \geq d(x,A)$

$$sech^2 d(x,B) \leq sech^2 d(x,A)$$

$$|A| \leq \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,A)dx.$$

For the opposite inequality, choose a sequence (A_k) of finite subsets of A converging to A in the Hausdorff metric (refer to definition 3.1).

Since $A_k \to A$ then $d(x,A_k) \to d(x,A)$ and since $sech^2$ is a decreasing function on $[0,\infty)$ then $0 \leq sech^2 d(x,A_k) \leq sech^2 d(x,A)$ for all x and k, then

$$lim_{k\to\infty}\int_{\mathbb{R}} sech^2 d(x,A_k)dx = \int_{\mathbb{R}} sech^2 d(x,A)dx$$

by dominated convergence theorem. And thus

$$|A| = \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,A)dx.$$

□

Example 3.8. *The magnitude of* $A = [0,1]$.

We know from section 3.3.1 that $|A| = 1 + \frac{1}{2} = \frac{3}{2}$. *Now we need to verify this result using proposition 3.7:*

$$
\begin{aligned}
|A| &= \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,A)dx \\
&= \frac{1}{2}\Big[\int_{-\infty}^{0} sech^2(-x)dx + \int_0^1 sech^2(0)dx + \int_1^{\infty} sech^2(x-1)dx\Big] \\
&= \frac{1}{2}\Big[-tanh(-x)\Big|_{-\infty}^{0} + x\Big|_0^1 + tanh(x-1)\Big|_1^{\infty}\Big] \\
&= \frac{1}{2}(1+2) \\
&= \frac{3}{2}.
\end{aligned}
$$

Example 3.9. *The effect of adding a singleton to a set.*

Let $A = [0,1] \cup \{2\}$.

A is a compact subset of $\mathbb{R}$ and therefore $|A| = \frac{1}{2}\int_{\mathbb{R}} sech^2 d(x,A)dx$.

$$
d(x,A) = \begin{cases} -x & \text{if } x < 0 \\ 0 & \text{if } 0 \le x < 1 \\ x-1 & \text{if } 1 \le x < \frac{3}{2} \\ 2-x & \text{if } \frac{3}{2} \le x \le 2 \\ x-2 & \text{if } x > 2. \end{cases}
$$

$$
\begin{aligned}
|A| = \frac{1}{2}\Big[\int_{-\infty}^{0} sech^2(-x)dx &+ \int_0^1 sech^2(0)dx + \int_1^{\frac{3}{2}} sech^2(x-1)dx + \int_{\frac{3}{2}}^{2} sech^2(2-x)dx \\
&+ \int_2^{\infty} sech^2(x-2)dx\Big]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2}\Big[1 + 1 + 2tanh\frac{1}{2} + 1\Big] \\
&= \frac{3}{2} + tanh\frac{1}{2}
\end{aligned}
$$

3.2.2 The magnitude of the cantor ternary set

In this section, we illustrate the work done in [4] to get the magnitude of the cantor ternary set. We need to construct the cantor set L_l for every l by considering a sequence converging to it. We start the construction by letting L_l^0 be the first approximation consisting of the two endpoints which are at a distance l.

Let ψ_1 and ψ_2 be two scalings of the real line such that each has a fixed point which is an endpoint of L_l^0 and it divides the segment into 3 segments of equal length.

Define $L_l^k = \psi_1(L_l^{k-1}) \cup \psi_2(L_l^{k-1})$ for $k \ge 1$.

The cantor set is $L_l = \cup_k L_l^k$, this is proved in work of Hutchinson (see[2]).

Now we can start by finding the magnitude of the finite approximations to the Cantor set and later deduce the magnitude of the Cantor set itself as done in [4].

Theorem 3.10. *The magnitude of the kth approximation to the Cantor set of length l is*

$$|L_l^k| = 1 + 2^k tanh\Big(\frac{l}{2.3^k}\Big) + \frac{1}{2}\sum_{i=1}^{k} 2^i tanh\Big(\frac{l}{2.3^i}\Big).$$

Proof. From section 3.1.1, if $A \subset \mathbb{R}$, then $|A| = 1 + \sum_{i=1}^{n} tanh(\frac{d_i}{2})$. The $|L_l^{k+1}|$ approximation consists of two copies of $|L_{l/3}^k|$ with distance $\frac{l}{3}$ apart. The pair of points in $|L_l^{k+1}|$ are either both on the same copy of $|L_{l/3}^k|$ or on two different copies and are $\frac{l}{3}$ apart[4].

We perform the proof by mathematical induction. We know that, $|L_l^0| = 1 + tanh(\frac{l}{2})$.

Suppose the equation is true for k and then

$L_l^{k+1} = \psi_1(L_{l/3}^k) \cup \psi_2(L_{l/3}^k)$

$$|L_{l/3}^k| = 1 + 1 + 2^k tanh\Big(\frac{l/3}{2.3^k}\Big) + \frac{1}{2}\sum_{i=1}^{k} 2^i tanh\Big(\frac{l/3}{2.3^i}\Big)$$

$$\begin{aligned}|L_l^{k+1}| &= 1 + 2^{k+1} tanh\Big(\frac{l}{2.3^{k+1}}\Big) + \frac{1}{2}\sum_{i=1}^{k} 2^{i+1} tanh\Big(\frac{l}{2.3^{i+1}}\Big)\\ &= 1 + 2^{k+1} tanh\Big(\frac{l}{2.3^{k+1}}\Big) + \frac{1}{2}\sum_{i=1}^{k+1} 2^i tanh\Big(\frac{l}{2.3^i}\Big).\end{aligned}$$

□

Theorem 3.11. *The magnitude of L_l, the Cantor set of length l is*

$$|L_l| = 1 + \frac{1}{2}\sum_{i=1}^{\infty} 2^i tanh\Big(\frac{l}{2.3^i}\Big).$$

Proof.

$$|L_l| = \lim_{k\to\infty} 1 + 2^k tanh\Big(\frac{l}{2.3^k}\Big) + \frac{1}{2}\sum_{i=1}^{k} 2^i tanh\Big(\frac{l}{2.3^i}\Big).$$

But $|tanh(x)| \leq x$ for all $x \geq 0$.

$$|tanh(\frac{l}{2.3^k})| \leq \frac{l}{2.3^k}.$$

Therefore, $\lim_{k\to\infty} tanh(\frac{l}{2.3^k}) = 0$

and

$$\sum_{i=1}^{\infty} 2^i tanh\Big(\frac{l}{2.3^i}\Big) \leq \sum_{i=1}^{\infty} \frac{2^i}{2.3^i} = \frac{l}{2}\sum_{i=1}^{\infty}\Big(\frac{2}{3}\Big)^i$$

which is a converging geometric series.

$$\sum_{i=1}^{\infty} 2^i tanh\Big(\frac{l}{2.3^i}\Big)$$

exists.

$$|L_l| = 1 + \frac{1}{2}\sum_{i=1}^{\infty} 2^i tanh\Big(\frac{l}{2.3^i}\Big).$$

$\square$

3.2.3 The magnitude of a circle

In this section, we calculate the magnitude of the circle as a subspace of Euclidean space.

Proposition 3.12. *Let C_l be a circle with circumference l, then the magnitude of C_l is given by*

$$|C_l| = \Big(\int_0^1 exp(\frac{-l}{\pi} sin(\pi s))ds\Big)^{-1}.$$

Proof. Consider, a circle of circumference l as a subset of $\mathbb{R}^2$. So the distance between the points x_1 and x_2 is given by $d(x_1, x_2) = \frac{l}{\pi} sin(\frac{\theta}{2})$.

Now, we approximate the circumference with a finite set of equidistant points. Define C_l^k to be the set of k points equally spaced around the circumference of the circle. This approximation is valid since given $\epsilon > 0$ and we choose k large enough so that the k points are distributed regularly over the unit circle have the property that the set $\{\epsilon$ -neighborhoods of the points$\}$ covers the circle.

This approximation yields a homogeneous finite metric space, therefore we can apply Speyer's formula (Proposition 2.7) to get the magnitude. So, for any point x in C_l^k:

$$|C_l^k| = \frac{n}{\sum_{x' \in C_l^k} e^{-d(x,x')}} = \frac{1}{\sum_{x' \in C_l^k} \frac{1}{n} exp(\frac{-l}{\pi} sin(\frac{\pi j}{n}))}$$

Now taking the limit as $k \to \infty$, we notice that $\sum_{x' \in C_l^k} \frac{1}{n} exp(\frac{-l}{\pi} sin(\frac{\pi j}{n}))$ is just a Riemann sum that has as limit an integral.

So as $k \to \infty$, $|C_l^k| \to \frac{1}{\int_0^1 exp(\frac{-l}{\pi} sin(\pi s))ds}$.

$|C_l^k| \to |C_l|$ since C_l is a compact positive definite space and C_l^k is a sequence of compact subspaces converging to C_l.

$|C_l| = \Big(\int_0^1 exp(\frac{-l}{\pi} sin(\pi s))ds\Big)^{-1}$ $\square$

3.3 Third Approach

In this section we present the generalization of the definition of magnitude using measures as done in [7].

Definition 3.13. *If X is a metric space, then a **weight measure** on X is a finite signed Borel measure ν on X such that for all $y \in X$*

$$\int_{x \in X} e^{-d(x,y)} d\nu(x) = 1.$$

If a weight measure ν exists, then

$$|X| = \int_X d\nu$$

$|X|$ is independent of ν.

Now, we will consider the spaces whose magnitude has already been calculated and re-calculated using the weight measure.

3.3.1 The magnitude of a segment in $\mathbb{R}$

In this section, we will consider the magnitude of a line segment of length l.

Proposition 3.14. *A weight measure on a line segment L_l of length l has weight measure $\nu = \frac{1}{2}(\delta_a + \delta_b + \mu)$ where μ is the Lebesgue measure on the line segment and δ_a and δ_b are the Dirac delta measures at the endpoints. And hence the magnitude of L_l is given by: $|L_l| = 1 + l/2$ [7].*

Proof. First, we check that the suggested measure is a weight measure.

$$\begin{aligned}
\int_{x \in L_l} e^{-d(x,y)} d\nu &= \int_{x \in L_l} e^{-d(x,y)} . \frac{1}{2}(d\delta_a + d\delta_b + d\mu) \\
&= \frac{1}{2}\Big(e^{-d(a,y)} + e^{-d(b,y)} + \int_a^b e^{-d(x,y)} dx\Big) \\
&= \frac{1}{2}\Big(e^{-(y-a)} + e^{-(d-y)} + \int_a^y e^{-(y-x)} dx + \int_y^b e^{-(x-y)} dx\Big) \\
&= 1.
\end{aligned}$$

Since we have a weight measure, we can proceed with calculating the magnitude.

$$|L_l| = \int_{L_l} d\nu = 1 + \frac{l}{2}.$$

□

3.3.2 The magnitude of the cantor ternary set

In order to study the magnitude of the cantor set, we will consider the magnitude of $X \backslash_{(a,b)}$ first.

Proposition 3.15. *Consider $X \subset \mathbb{R}$ and $[a,b] \subset X$ with ν a weight measure on X and $\nu\Big|_{[a,b]} = \frac{1}{2}\mu\Big|_{[a,b]}$. The weight measure on $X\backslash_{(a,b)}$ is given by*

$$\nu\Big|_{X\backslash_{(a,b)}} + \frac{1}{2}tanh\Big(\frac{b-a}{2}\Big)\Big(\delta_a + \delta_b\Big).$$

And the magnitude is given by:

$$|X\backslash_{(a,b)}| = |X| - \frac{b-a}{2} + tanh(\frac{b-a}{2}).$$

Proof. First, we'll check that the suggested measure is a weight measure.

$$\begin{aligned}
&\int_{x\in X\backslash_{(a,b)}} e^{-d(x,y)} d\Big(\nu + \frac{1}{2}tanh\Big(\frac{b-a}{2}\Big)\Big(\delta_a + \delta_b\Big)\Big) \\
&= \int_{x\in X\backslash_{(a,b)}} e^{-d(x,y)} d\nu + \frac{1}{2}\int_{x\in X\backslash_{(a,b)}} e^{-d(x,y)} d\Big(tanh\Big(\frac{b-a}{2}\Big)\Big(\delta_a + \delta_b\Big)\Big) \\
&= \int_{x\in X} e^{-d(x,y)} d\nu - \int_{x\in[a,b]} e^{-d(x,y)} d\nu\Big|_{[a,b]} + \frac{1}{2}\int_{x\in X\backslash_{(a,b)}} e^{-d(x,y)} tanh\Big(\frac{b-a}{2}\Big)\Big(d\delta_a + d\delta_b\Big) \\
&= 1 - \int_a^b e^{-(x-y)}\frac{1}{2}d\mu + \frac{1}{2}tanh\Big(\frac{b-a}{2}\Big)\Big(e^{y-a} + e^{y-b}\Big) \\
&= 1 + \frac{1}{2}e^{y-b} - \frac{1}{2}e^{y-a} + \frac{1}{2}e^{y-a} - \frac{1}{2}e^{y-b} \\
&= 1.
\end{aligned}$$

Now we can calculate the expression for the magnitude:

$$\begin{aligned}
|X\backslash_{(a,b)}| &= \int_{x\in X\backslash_{(a,b)}} \Big(d\nu\Big|_{(X\backslash_{(a,b)})} + \frac{1}{2}tanh(\frac{b-a}{2})\Big(d\delta_a + d\delta_b\Big)\Big) \\
&= \int_{x\in X} d\nu - \int_{x\in[a,b]} d\nu + \frac{1}{2}tanh(\frac{b-a}{2})\int_{x\in X\backslash_{(a,b)}} \Big(d\delta_a + d\delta_b\Big) \\
&= |X| - \frac{1}{2}\int_{x\in[a,b]} d\mu + \frac{1}{2}tanh(\frac{b-a}{2})(1+1) \\
&= |X| - \frac{b-a}{2} + tanh(\frac{b-a}{2}).
\end{aligned}$$

□

Using the previous result to find the magnitude of a cantor set which will be approximated with subsets of $[0,l]$ of non-zero measure[7].

Proposition 3.16. *The cantor ternary set L_l of length l has magnitude given by:*

$$|L_l| = 1 + \sum_{i=1}^{\infty} 2^{i-1} tanh(\frac{l}{2\times 3^i})$$

Proof. Define L_j^0 to be the line segment of length l. As in section 3.3.1, L_j^0 has weight measure $\nu_l^0 = \frac{1}{2}(\mu + \delta_0 + \delta_l)$. Now, inductively we define L_l^j and its weight measureν_l^j by removing from

L_l^{j-1} the open middle third of each maximal subinterval. The length of each removed open interval is $(1/3)^j l$.

We construct the weight measure for this sequence, using proposition 3.9. Let S_i be the set of the newly exposed endpoints at the j^{th} level. Then the weight measure is given by:

$$\nu_l^j = \Big(\mu\Big|_{L_l^j} + \delta_0 + \delta_l + \sum_{i=1}^{j} tanh(\frac{l}{2 \times 3^i}) \sum_{s \in S_i} \delta_s\Big)$$

.

Consider the sequence $\{\nu_l^j\}_{j=1}^{\infty}$ on the interval $[0,l]$. In the Banach space of finite signed Borel measures on $[0,l]$, this sequence converges in the total variation norm $\nu_l^j \to \nu_l$ where

$$\nu_l = \Big(\delta_0 + \delta_l + \sum_{i=1}^{\infty} tanh(\frac{l}{2 \times 3^i}) \sum_{s \in S_i} \delta_s\Big).$$

Note that, a sequence ν_l^j of measures defined on the same measure space is said to converge to a measure ν_l in total variation distance if for every $\epsilon > 0$ there exists an N such that for all $n > N$: $\|\nu_l^j - \nu_l\|_{TV} < \epsilon$ and this is true since $\mu\Big|_{Cantorset} = 0$.

Since convergence in total variation norm implies pointwise convergence as well on any continuous function on $[0,l]$ such as $e^{-d(.,y)}$ for y $\in T_l$ hence $\displaystyle\int_{x \in [0,l]} e^{-d(x,y)} d\nu_l^j \to \int_{x \in [0,l]} e^{-d(x,y)} d\nu$.[7]

But as $y \in L_l \subset L_l^j$ for all j and ν_l^j is a weight measure, then each of the integrals on the left hand side is 1 and therefore $\displaystyle\int_{x \in [0,l]} e^{-d(x,y)} d\nu = 1$ and hence ν is a weight measure on T_l.

Now we can calculate the magnitude of L_l:

$$\begin{aligned}
|L_l| &= \int_{L_l} d\nu_l \\
&= \frac{1}{2} \int_{L_l} \Big(d\delta_0 + d\delta_l + \sum_{i=1}^{\infty} tanh\Big(\frac{l}{2 \times 3^i}\Big) \sum_{s \in S_i} d\delta_s\Big) \\
&= 1 + \frac{1}{2} \int_{L_l} \sum_{i=1}^{\infty} tanh\Big(\frac{l}{2 \times 3^i}\Big).2^i \\
&= 1 + \sum_{i=1}^{\infty} tanh(\frac{l}{2 \times 3^i}).2^{i-1}.
\end{aligned}$$

□

3.3.3 The magnitude of a circle

Proposition 3.17. *The magnitude of a circle C_l of circumference l is given by*

$$|C_l| = \frac{l}{\displaystyle\int_0^{2\pi} e^{-\frac{\pi}{l} sin\frac{\theta}{2}} d\theta}.$$

Proof. Theorem1 in [7] states that if X is a homogeneous metric space and μ is an invariant measure on X then $\displaystyle\int_{x\in X} e^{-d(x,y)}d\mu$ is independent of y. If this quantity is non-zero and finite then a weight measure on X is given by

$$\frac{\mu}{\displaystyle\int_{x\in X} e^{-d(x,y)}d\mu}.$$

Thus the magnitude is given by

$$|X| = \frac{\displaystyle\int_{x\in X}\mu}{\int_{x\in X} e^{-d(x,y)}d\mu}.$$

$$|C_l| = \frac{\displaystyle\int_{x\in C_l}\mu}{\displaystyle\int_{x\in C_l} e^{-d(x,y)}d\mu} = \frac{l}{\displaystyle\int_0^{2\pi} e^{-\frac{\pi}{l}sin\frac{\theta}{2}}d\theta}$$

□

Chapter 4

Dimensions

In this chapter we present the proof that for compact metric spaces, the diversity dimension and the Minkowski dimension are equal. Then we use a relation between diversity function and magnitude function for subsets of Euclidean subspaces to show that diversity dimension and magnitude dimension are equal for Euclidean subspaces and hence in this case the magnitude and Mikowski dimension are equal. This part is based on the work done by Mark Meckes in [6].

4.1 Magnitude Dimension

The magnitude dimension of a compact space as described in [3] is the growth of the magnitude function. We will start by defining the **upper magnitude dimension** of a compact metric space A to be:

$$\overline{dim}_{Mag}A := \limsup_{t\to\infty} \frac{log|tA|}{logt}$$

and the **lower magnitude dimension** of A is:

$$\underline{dim}_{Mag}A := \liminf_{t\to\infty} \frac{log|tA|}{logt}.$$

The magnitude dimension is defined when the above two limits are equal and thus

$$dim_{Mag}A = \lim_{t\to\infty} \frac{log|tA|}{logt}$$

if it exists.

4.2 Minkowski Dimension

In order to define the Minkowski dimension, we will need to define two numbers as suggested in [6]:

- the **packing number** $M(A, \epsilon)$ is the minimum number of disjoint closed ϵ-balls is A.
- the **covering number** $N(A, \epsilon)$ is the minimum number of of closed ϵ-balls needed to cover A.

Now we can define the **upper and lower Minkowski dimension** of A to be:

$$\overline{dim}_{Mink}A := \limsup_{\epsilon\to 0^+}\frac{logN(A,\epsilon)}{log(1/\epsilon)} = \limsup_{\epsilon\to 0^+}\frac{logM(A,\epsilon)}{log(1/\epsilon)}, \underline{dim}_{Mink}A := \liminf_{\epsilon\to 0^+}\frac{logN(A,\epsilon)}{log(1/\epsilon)} = \liminf_{\epsilon\to 0^+}\frac{logM(A,\epsilon)}{log(1/\epsilon)}.$$

If $\overline{dim}_{Mink}A = \underline{dim}_{Mink}A$, then $\dim_{Mink} A = \lim_{\epsilon\to 0^+}\dfrac{logM(A,\epsilon)}{log(1/\epsilon)}$.

4.3 Diversity Dimension

As defined in [6], for a compact metric space A, the **maximum diversity** of A is given by

$$|A|_+ = \sup_{\mu\in P(A)}\Big(\int\int e^{-d(a,b)}d\mu(a)d\mu(b)\Big)^{-1}$$

where $P(A)$ is the space of Borel probability measure on A.

Now, we define the **upper and lower diversity dimension** of A to be:

$$\overline{dim}_{Div}A := \limsup_{t\to\infty}\frac{log|tA|_+}{logt}, \underline{dim}_{Div}A := \liminf_{t\to\infty}\frac{log|tA|_+}{logt}$$

and

$$dim_{Div}A = \lim_{t\to\infty}\frac{log|tA|_+}{logt}$$

when it exists.

We have from (4.4) in [6]

$$|A|_+ \leq |A|$$

for any compact PDMS A.

Then

$$\underline{dim}_{Div}A \leq \underline{dim}_{Mag}A; \overline{dim}_{Div}A \leq \overline{dim}_{Mag}A$$

4.4 Relationship Between Diversity, Minkowski and Magnitude Dimensions

Theorem 4.1. *For any compact metric space A,*

$$\underline{dim}_{Div}A = \underline{dim}_{Mink}A \text{ and } \overline{dim}_{Div}A = \overline{dim}_{Mink}A.$$

Thus, $dim_{Div}A$ is defined if and only if $dim_{Mink}A$ is defined, and in that case $dim_{Div}A$=$dim_{Mink}A$.

Proof. First, we show that $\underline{dim}_{Div}A \leq \underline{dim}_{Mink}A$.

Let $\epsilon > 0$, $t > 0$ and $\mu \in P(A)$ be given, then for each a $\in$ A

$\displaystyle\int e^{-td(a,b)}d\mu(b) \geq \int_{B(a,\epsilon)} e^{-td(a,b)}d\mu(b) \geq e^{-t\epsilon}\mu(B(a,\epsilon))$

$\displaystyle\int\int e^{-td(a,b)}d\mu(a)d\mu(b) \geq \int e^{-t\epsilon}\mu(B(a,\epsilon))d\mu(a)$

$\Big(\int\int e^{-td(a,b)}d\mu(a)d\mu(b)\Big)^{-1} \leq e^{t\epsilon}\Big(\int \mu(B(a,\epsilon))d\mu(a)\Big)^{-1}$

Here we can use Jensen's inequality to get:

$\Big(\int \mu(B(a,\epsilon))d\mu(a)\Big)^{-1} \leq \int \frac{1}{\mu(B(a,\epsilon))}d\mu(a).$

Let N be the covering number $N(A,\epsilon/2)$, and let $a_1,...,a_N \in A$ be the centers of the closed $\frac{\epsilon}{2}$-balls needed to cover A.

If $a \in B(a_i,\epsilon/2)$ then $B(a_i,\epsilon/2) \subseteq B(a,\epsilon)$

$$\begin{aligned}\int \frac{1}{\mu(B(a,\epsilon))}d\mu(a) &\leq \sum_{i=1}^{N}\int\limits_{B(a_i,\epsilon/2)} \frac{1}{\mu(B(a,\epsilon))}d\mu(a)\\ &\leq \sum_{i=1}^{N}\int\limits_{B(a_i,\epsilon/2)} \frac{1}{\mu(B(a_i,\epsilon/2))}d\mu(a)\\ &= \sum_{i=1}^{N}\frac{\mu(B(a_i,\epsilon/2))}{\mu(B(a_i,\epsilon/2))}\\ &= N.\end{aligned}$$

Note that the above summation is taken over all the $i's$ where $\mu(B(a_i,\epsilon/2)) > 0$.

So

$$|tA|_+ = \sup_{\mu\in P(A)}\Big(\int\int e^{-td(a,b)}d\mu(a)d\mu(b)\Big)^{-1} \leq e^{t\epsilon}N(A,\epsilon/2),$$

and letting $\epsilon = 2/t$

$$|tA|_+ \leq e^2N(A,1/t) = log|tA|_+ \leq log(e^2N(A,1/t))$$

Then taking logs and dividing by $logt$, and passing to the limit as t tends to infinity, we get $\underline{dim}_{Div}A \leq \underline{dim}_{Mink}A$ and $\overline{dim}_{Div}A \leq \overline{dim}_{Mink}A$.

We still need to show the other inequality $\underline{dim}_{Mink}A \leq \underline{dim}_{Div}A$ and $\overline{dim}_{Div}A \leq \overline{dim}_{Mink}A$.

To get the opposite inequality, we can use the packing number. So let M be the packing number $M(A,\epsilon)$ and $a_1,...a_M$ be the centers of disjoint closed ϵ-balls in A.

Define the following measure:

$$\mu = \frac{1}{M}\sum_{i=1}^{M}\delta_{a_i} \in P(A).$$

Since the balls are disjoint, then for each $a \in A$, there is at most one a_i in the ball $B(a,\epsilon)$, so

$$e^{-td(a,b)}d\mu(b) = \frac{1}{M}\sum_{i=1}^{M}e^{-td(a,a_i)} \leq \frac{1}{M}(e^0+(M-1)e^{-t\epsilon}) \leq \frac{1}{M}+e^{-t\epsilon}$$

$$\int\int e^{-td(a,b)}d\mu(b)d\mu(a) \leq \int \frac{1}{M}+e^{-t\epsilon}d\mu$$

$$\frac{1}{|tA|_+} \leq \Big(\frac{1}{M} + e^{-t\epsilon}\Big) \times \frac{1}{M}\sum_{i=1}^{M}\delta_{a_i} = \frac{1}{M} + e^{-t\epsilon}.$$

As done in [6], let $\epsilon(t) = \dfrac{log(2|tA|_+}{t}$ for $t \geq 1$.

Then $e^{-t\epsilon} = e^{-t(\frac{log(2|tA|_+}{t})} = \dfrac{1}{2|tA|_+}$

$\dfrac{1}{|tA|_+} \leq \dfrac{1}{M(A,\epsilon)} + \dfrac{1}{2|tA|_+}$

$\dfrac{1}{M(A,\epsilon(t))} \geq \dfrac{1}{2|tA|_+}$

$M(A,\epsilon(t)) \leq 2|tA|_+$

Now, $\epsilon(t)$ is continuous and strictly decreasing. If $\epsilon(t)$ is bounded below by a positive constant then $\underline{dim}_{Div}A = \infty$ and we get the desired result [6].

Suppose $\epsilon(t) \to 0$ as $t \to \infty$,

$$\begin{aligned}\frac{M(A,\epsilon(t))}{log(1/\epsilon(t))} \leq \frac{log(2|tA|_+)}{log(1/\epsilon(t))} &= \frac{logt}{logt}.\frac{log(2|tA|_+)}{log(1/\epsilon(t))}\\ &= \frac{log(2|tA|_+)}{logt}.\frac{logt}{log(1/\epsilon(t))}\\ &= \frac{log(|tA|_+)}{logt}.\frac{logt}{log(1/\epsilon(t))} + o(1)\end{aligned}$$

So,

$$\frac{log(2|tA|_+)}{log(1/\epsilon(t))} = \frac{log(|tA|_+)}{logt}.\frac{logt}{log(1/\epsilon(t))}$$

$$\frac{logt}{log(1/\epsilon(t))} = \frac{log(\frac{t}{log(|tA|_+)})}{logt} = 1 - \frac{loglog(2|tA|_+)}{logt}$$

If $t \to \infty$ and $\dfrac{|tA|_+}{logt}$is bounded above so $\dfrac{logt}{log(1/\epsilon(t))} \to 1$

$$\frac{M(A,\epsilon(t))}{log(1/\epsilon(t))} \leq \frac{log(|tA|_+)}{logt}(1+o(1)).$$

So

$$\underline{dim}_{Mink}A \leq \liminf_{t\to\infty}\frac{M(A,\epsilon(t))}{log(1/\epsilon(t))} \leq \liminf_{t\to\infty}\frac{log(|tA|_+)}{logt} = \underline{dim}_{Div}A.$$

Similarly, $\overline{dim}_{Mink}A < \overline{dim}_{Div}A$. □

Corollary 4.2. *If A is a compact subset of $\mathbb{R}^n$, then $\overline{dim}_{Mag}A = \overline{dim}_{Mink}A$ and $\underline{dim}_{Mag}A = \underline{dim}_{Mink}A$. So, $dim_{Mag}A$ is defined if and only if $dim_{Mink}A$ is defined and thus $dim_{Mag}A$=$dim_{Mink}A$.*

Proof. From [6], we have the following results for each compact $A \subseteq \mathbb{R}^n$ and for each positive integer n there exits $k_n > 0$ such that:

$$|tA|_+ \leq |tA| \leq k_n |tA|_+$$

where A is a compact subset of $\mathbb{R}^n$.

Combining this result with theorem 4.1, we get that for A being a compact subset of $\mathbb{R}^n$

$$\overline{dim}_{Mag} A = \overline{dim}_{Div} A \text{ and } \underline{dim}_{Mag} A = \underline{dim}_{Div} A$$
$$\overline{dim}_{Mag} A = \overline{dim}_{Mink} A \text{ and } \underline{dim}_{Mag} A = \underline{dim}_{Mink} A \qquad \square$$

The relationship between the magnitude dimension and Minkowski dimension for subset of Euclidean spaces also holds true for compact homogeneous spaces as shown by Mark Meckes in [6].

Bibliography

[1] F. Abi-Khuzam, *The magnitude of a segment* unpublished notes

[2] J. E. Hutchinson, *Fractals and self-similarity*, Indiana University Mathematics Journal 30 (1981), no. 5, 713747.

[3] T. Leinster. *The magnitude of metric spaces.* Doc. Math., 18:857905, 2013.

[4] T. Leinster and S. Willerton. *On the asymptotic magnitude of subsets of Euclidean space.* Geom. Dedicata, 164(1):287310, 2013.

[5] M. Meckes, *Positive definite metric spaces* arXiv preprint. http://arxiv.org/abs/1012.5863v3

[6] M. Meckes, *Magnitude, diversity, capacities, and dimensions of metric spaces* arXiv preprint. http://arxiv.org/abs/1308.5407v2

[7] S. Willerton, *On the magnitude of spheres, surfaces and other homogeneous spaces*, arXiv preprint. http://arxiv.org/abs/1005.4041v1