FINITE METRIC SPACES AND THEIR EMBEDDING INTO LEBESGUE SPACES

RYAN HOPKINS

Abstract. The properties of the metric topology on infinite and finite sets are analyzed. We answer whether finite metric spaces hold interest in algebraic topology, and how this result is generalized to pseudometric spaces through the Kolmogorov quotient. Embedding into Lebesgue spaces is analyzed, with special attention for Hilbert spaces, ℓ^p, and $\mathbb{E}^N$.

Contents

1. Introduction

Finite metric spaces are simple objects, a finite collection of points with a real distance defined between each pair. Despite their apparent simplicity, they are intriguing. From the perspective of algebraic topology, they have no interest as discrete spaces. Although relaxing metrics to pseudometrics appears to provide finite metric spaces with more interest, pseudometric spaces are homotopically equivalent to the discrete space formed when they are passed through the Kolmogorov quotient. Despite their uninteresting topogical structure, finite metric spaces have applications to computer science. Many physical systems can be modeled with finite points and distances between them, so computer scientists are motivated to embed finite metric spaces into host spaces like $\mathbb{R}^N$ where detailed analysis can be

Date: August 17 2015.

done. Perfect embeddings cannot always be achieved, so the study of the distortion needed for embeddings and when isometric embeddings exist is a rich area.

This paper first considers finite metric spaces from a topological perspective, highlighting general properties and showing why they seem to hold no interest topologically. The last section surveys the literature on embeddings of finite metric spaces.

2. Finite Metric Spaces

Finite spaces have different metrization and pseudometrization conditions and their metrics can be represented in convenient ways.

2.1. Pseudometrizing Finite Spaces.

Definition 2.1. A *pseudometric* is a function $d : X \times X \to \mathbb{R}$ which satisfies the following properties:

i. $d(x, x) = 0 \ \forall x \in X$
ii. $d(x, y) \geq 0$
iii. $d(x, y) = d(y, x) \ \forall x, y \in X$
iv. $d(x, y) + d(y, z) \geq d(x, z) \ \forall x, y, z \in X$

This definition is a weakening of the standard metric. Two distinct points may have a distance of zero. Pseudometrics are sometimes referred to as *semimetrics*.

Definition 2.2. A space X is *pseudometrizable* if there is a pseudometric d on X that induces the topology of X.

Definition 2.3. A space is R_0 if each pair of topologically distinct points (i.e. points which do not have the same set of neighborhoods) have some neighborhood not containing the other point.

Theorem 2.4. *A finite topological space is pseudometrizable iff it is R_0.*

Proof. Given a topological space X and points x and y in X, define x $\equiv$ y to mean that x and y are topologically indistinguishable.

Define the *standard discrete pseudometric* to be:

$$d(x, y) = \begin{cases} 0 & \text{if x} \equiv \text{y} \\ 1 & \text{if x} \not\equiv \text{y} \end{cases}$$

Given x $\not\equiv$ y, take neighborhoods B(x,($\frac{1}{2}$)) and B(y,($\frac{1}{2}$)) of x and y so that

$$B(x, (\frac{1}{2})) \cap B(y, (\frac{1}{2})) = \emptyset$$

This metric induces a topology on X where every topologically distinguishable pair is separated.

If a finite space is R_0 with its given topology, then it can be given this topology which separates topologically distinguishable points, satisfying the R_0 condition as well as inducing a topology which puts families of points equivalent to the given topology into the same neighborhoods.

Take a space X to be pseudometrizable. Then its metric topology forms open balls around topologically distinguishable points which can be separated.

If no points in the space have distinct neighborhoods (i.e. the pseudometric outputs 0 given any two points), then there are no topologically distinguishable points, so the space is vacuously R_0. □

2.2. **Representing Metrics on Finite Spaces.**

A metric on a finite space can be explicitly defined by $\binom{n}{2}$ non-negative numbers, where each number corresponds to a distance between two points. This property of finite metric spaces allows them to represented in convenient ways, most importantly with matrices and graphs.

2.2.1. *Matrix Representation.*

Take a finite metric space (X,d) with points (x_0,x_1,...,x_n). Construct an n × n matrix with entries ($a_{i,j}$) giving the distance between point i and point j in the space. Then the following characteristics can be observed.

1. $d(x_i, x_j) \geq 0$ for all $0 \leq i, j \leq n$ so the matrix is comprised of nonnegative real numbers.
2. $d(x_i, x_i) = 0$ for all $0 \leq i \leq n$ so the diagonal of the matrix is 0.
3. $d(x_i, x_j) = d(x_j, x_i)$ for all $0 \leq i, j \leq n$ so the matrix equals its transpose.

Thus any finite metric space has a real, positive, symmetric matrix containing all the information of its metric.

2.2.2. *Graph Representation.*

The matrix defined by the finite metric space can be translated to an undirected, no loop, weighted, finite graph. Given a finite metric space (X,d) with points (x_0,x_1,...,x_n), a graph G with n vertices and $\binom{n}{2}$ weighted edges giving the distance between vertices can be constructed to represent it.

The distance function defines a distance between any two points of the space, so each vertex of the graph connects to every other vertex, forming a complete graph.

Metrics satisfy the triangle inequality, so all edges may not be necessary if the *shortest path metric* is used on the graph.

Definition 2.5. Given a weighted graph G, the *shortest path metric* is a metric which defines the distance between two vertices to be the length of the shortest path between them. If the two vertices are not connected, the distance is said to be infinite.

Theorem 2.6. *A graph G with n vertices and the shortest path metric represents an n point finite metric space (X,d) iff it is undirected, no loop, weighted and connected.*

Proof. Set each vertex in G to represent a distinct point in the underlying set X.

The properties of a metric give rise to the conditions necessary for the graph.

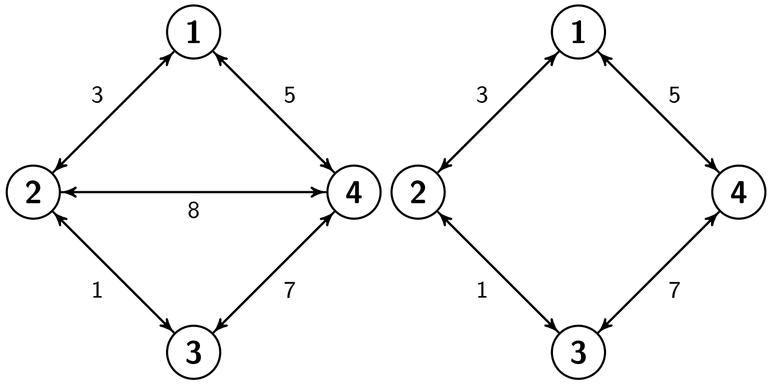
1. $d(x_i, x_j) = d(x_j, x_i)\ \forall\ 0 \leq i, j \leq n$ G must be undirected
2. $d(x_i, x_i) = 0\ \forall\ 0 \leq i \leq n$ G must have no loops
3. $d(x_i, x_j) \geq 0\ \forall\ 0 \leq i \leq n$ G must be weighted with nonnegative real values
4. $d(x_i, x_j) < \infty\ \forall\ 0 \leq i, j \leq n$ G must be connected

The triangle inequality means that the shortest path metric must be used.

Conversely, a graph fulfilling the above properties can be made into a finite metric space if the vertices are made into the underlying set and the shortest path metric is made into the metric on that set. □

Definition 2.7. It may be possible to obtain a graph with fewer than $\binom{n}{2}$ edges (i.e. not a complete graph) to represent the finite metric space. When all edges which do not alter the output of the shortest path metric are dropped, the *critical graph* is obtained.

Example 2.8. Where the triangle inequality is satisfied by an equality an edge can be removed. In this case a critical graph is obtained.



3. The Problem with Finite Metric Spaces

Finite metric spaces are of no interest to algebraic topologists as they induce the discrete topology on the space. This section illustrates why this is the case and how an indiscrete pseudometric space can be made into a discrete space when it is made T_0 through the Kolmogorov Quotient.

3.1. **The Discrete Topology.**

Definition 3.1. The *discrete topology* is the finest topology possible on a set. Every subset is an open set (and therefore every subset is also a closed set). Every point separates the space in this topology, so it is called the discrete topology.

The fact that finite metric spaces have the discrete topology can be proved directly, or illustrated through Lipschitz equivalence of metrics.

Theorem 3.2. *Any metric on a finite space induces the discrete topology.*

Proof. Take a finite metric space (X,d). If every point in the space is open, then all of their possible unions are open, giving the discrete topology.

For any x $\in$ X, find r = $\min_{y \in X}$ (d(x,y)). This r exists and is nonzero as X is finite and d(x,y) $>$ 0 for x $\neq$ y. Then the open ball of radius r about x contains only x. The set $\{x\}$ is open. □

Theorem 3.3. *A finite space is metrizable iff it is discrete.*

Proof. Given a finite space with the discrete topology, the discrete metric ensures that every point is in a singleton open set (any open ball of radius less than 1) and so the finite space can be metrized.

Conversely, any finite space can be metrized in order to give the discrete topology. In fact, as proved above, the discrete topology is the only possible metric topology given to a finite space. □

3.2. **The Kolmogorov Quotient.**

Finite pseudometric spaces allow distinct points to have the same open neighborhoods in the induced topology. This seems to give them greater topological interest as they are not necessarily discrete. The Kolmogorov quotient provides a way to identify the topologically indistinguishable points and form a T_0 space. In this case, the T_0 space would be a metric space. Denote the Kolmogorov quotient of a space X by K(X).

Definition 3.4. A space is *T_0* if for every pair of distinct points, at least one of them has an open neighborhood not containing the other. In a T_0 space all points are topologically distinguishable.

Definition 3.5. There is a quotient space of any topological space which is always T_0. This is the *Kolmogorov Quotient*. This quotient space is formed under the equivalence relation which identifies points with the same open neighborhoods.

A pseudometric space is converted into a metric space through a Kolmogorov quotient by *metric identificaiton*.

3.2.1. *Metric Identification.*

Take (X,d) to be a pseudometric space with x, y $\in$ X. Set x $\sim$ y if d(x,y) = 0. Define $X^* = X/\sim$. Construct a metric d^* on X^* by setting $d^*([x], [y]) = d(x,y)$. Then (X^*, d^*) is a metric space.

Proposition 3.6. *Metric $d^*([x], [y]) = d(x,y)$ is well-defined*

Proof. If d^* is well-defined, then this equality will hold regardless of the choice of point in the equivalence class [x] and that d^* is a metric. It is clear that d^* is a metric as it inherits properties from metric d. Take $x_1, x_2 \in [x]$ and y $\in$ [y]. Then d^* is well-defined if $d^*(x_1,y) = d^*(x_2,y) = d(x,y)$. Take $d^*(x_1, y) = d(x,y)$. By the triangle inequality on d^*, $d^*(x_1,x_2) + d^*(x_2,y) \geq d^*(x_1,y)$. Because $x_1 \sim x_2$, $d^*(x_1,x_2) = 0$, so $d^*(x_2,y) = d^*(x_1,y)$. This means that d^* is well-defined as it does not depend on choice of representative from the equivalence class. □

Theorem 3.7. *Metric identification preserves the metric induced topology.*

Proof. To show that metric identification does preserve the topology of a pseudometric space (X,d) after passing to the quotient (X^*, d^*), it needs to be shown that the set A $\subset$ X is open iff set [A] (the set of all [x] where x is in A) is open in (X^*, d^*).

Take A $\subset$ (X,d), open. Then $\forall$ x $\in$ A, there is an open ball around x which is contained in A. Identify all x, y $\in$ such that d(x,y) = 0. These equivalence classes are made of points distance zero from each other, so the set of open balls $[B(x,\epsilon)]$ for a given [x], all overlap. □

3.2.2. *Kolmogorov Quotient of Pseudometric Spaces.*

Theorem 3.8. *The topology induced by metric identification forms a quotient space that is the Kolmogorov quotient.*

Proof. Take (X,d) a pseudometric space with identified metric as above.

To prove that this quotient space is a Kolmogorov quotient, it must be shown that the relation $\sim$ is an equivalence relation and that topology induced by d^* on $X/\sim$ forms K(X).

1. The Relation $\sim$ is an equivalence relation
   i. Reflexivity: d(x,x) = 0 $\forall$ x $\in$ X, so x $\sim$ x.
   ii. Symmetry: d(x,y) = d(y,x) $\forall$ x = y $\in$ X, so if d(x,y) = 0, then d(y,x) = 0 so if x $\sim$ y, then y $\sim$ x.

   iii. By the triangle inequality, d(x,y) + d(y,z) $\geq$ d(x,z) $\forall$ x,y,z $\in$ X. If x $\sim$ y and y $\sim$ z, then d(x,y) + d(y,z) = 0, so d(x,z) $\geq$ 0, so d(x,z) = 0.
2. The topology induced by d^* on X/$\sim$ forms K(X). For the topology induced by d^* on X/$\sim$ to be K(X), the equivalence classes must be comprised of topologically indistinguishable points. Take x,y $\in$ X, with x and y topologically distinguishable. Then there is an open subset U of X where x $\in$ U but y $\notin$ U. This means that there an open ball of some radius about x that does not contain y, so d(x,y) > 0, so x $\nsim$ y. Conversely, if x and y are topologically indistinguishable, then there is no open ball containing only one of the points. Then each B(x, $\frac{1}{n}$) must contain both x and y, so d(x,y) must be zero. This means that the topology induced by d^* on X/$\sim$ is putting only topologically indistinguishable points into equivalence classes. This, taken with Theorem 3.18 above, shows that this quotient forms K(X).

□

3.2.3. *Homotopy Equivalence of the Kolmogorov Quotient.*

Finite pseudometric spaces (in fact all finite spaces) are homotopy equivalent to their Kolmogorov Quotient K(X).

Definition 3.9. Take X,Y topological spaces and maps f: X $\to$ Y and g: X $\to$ Y. Maps f and g are *homotopic* if there is a continuous map h: X $\times$ [0,1] $\to$ Y where h(x,0) = f(x) $\forall$ x $\in$ X and h(x,1) = g(x) $\forall$ x $\in$ X. Denote this relation f $\simeq$ g.

Definition 3.10. Take X,Y topological spaces. Spaces X and Y are *homotopically equivalent* if there are continuous maps f: X $\to$ Y and g: Y $\to$ X where f $\circ$ g $\simeq$ Id_Y and g $\circ$ f $\simeq$ Id_X. Denote this relation X $\simeq$ Y.

Theorem 3.11. *Every finite space is homotopically equivalent to a T_0 space, K(X)* [3].

Corollary 3.12. *Any finite pseudometric space X is homotopically equivalent to its Kolmogorov Quotient, K(X), with K(X) being a finite metric space.*

4. Embedding Finite Metric Spaces

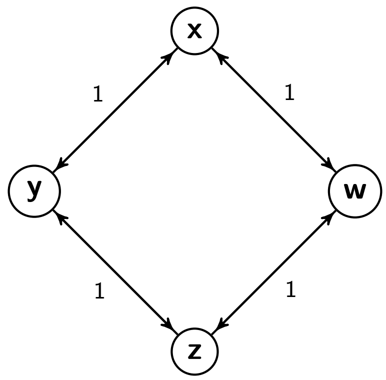
Despite the properties explored above, finite metric spaces are of interest to fields other than algebraic topology. In fields like microbiology, large tables of numbers are generated and need to be analyzed. It can be difficult to work with large tables, meaning that a representation in Euclidean space is desirable. An embedding would offer a way to see the distribution and behavior of the points of the metric space. In addition, a metric space with n points could be described in 2n numbers instead of $\binom{n}{2}$ numbers.

The interest in representing combinatorial objects like finite metric spaces in this way comes from a wider interest in the *geometrizaiton of combinatorial objects*, which is a method used to transform large amounts of information into a usable form.

Considering the equivalence between linear graphs and finite metric spaces given above, it would seem that all finite metric spaces could be represented in $\mathbb{R}^N$ for some finite N. This is not the case.

The distance metric on the weighted graph representing the finite metric space is the shortest path metric. In $\mathbb{R}^N$, the shortest path between two points is a straight line, so if equality holds in the triangle equality, those three points lie on the same line in $\mathbb{R}^N$. This fact will mean that not all finite metric spaces can be embedded without distorting the distances between points. This is illustrated in the following example.

Example 4.1. Take finite metric space (X,d) with 4 points represented by the weighted graph below with distance given by the shortest path metric.



This is a simple 4 cycle with edges of uniform length. Note that

$$d(x,z) = d(x,y) + d(y,z) = 2 \text{ and } d(x,z) = d(x,w) + d(w,z) = 2$$

This fact will give a contradiction when an embedding is done. Embed this metric space in $\mathbb{R}^N$. There are then two minimal paths between x and z and both obtain equality with the triangle inequality. As explained above, the fact that

$$d(x,z) = d(x,y) + d(y,z) \text{ and } d(x,z) = d(x,w) + d(w,z)$$

implies that points x,y,z are collinear, as are x,w,z. Line segments xyz and xwz are the same as they have the same endpoints. Because y and w are both distance 1 away from x on the same line, they are distance zero from each other. This implies that y = w, contradicting the fact that X has 4 points.

The graph must be *distorted* to be represented in $\mathbb{R}^N$.

Definition 4.2. Take metric spaces (X,d_X) and (Y,d_Y) and a function f: X $\rightarrow$ Y. Then the *distortion* of f can be realized by its Lipschitz constants. The *expansion* of f is defined as

$$\|f\|_{Lip} = \sup_{x,y \in X} \frac{d_Y(f(x), f(y))}{d_X(x,y)}$$

The *contraction* of f is given by

$$\|f\|_{Lip}^{-1} = \sup_{x,y \in X} \frac{d_X(x,y)}{d_Y(f(x), f(y))}$$

The *distortion* of f is given by

distortion(f) = contraction(f) * expansion(f) = $\|f\|_{Lip}^{-1} * \|f\|_{Lip}$

This is equivalent to finding the closest a, b $\in \mathbb{R}$ such that

$$a \geq \frac{d_Y(f(x), f(y))}{d_X(x,y)} \geq b$$

and defining distortion(f) = $\frac{a}{b}$.

Remark 4.3. A mapping f: X $\rightarrow$ Y is an *isometry* if $\frac{a}{b} = 1$, that is, all distances are preserved up to scaling.

Definition 4.4. Take metric spaces (X,d) and (Y,d$'$). Then (X,d) is *isometrically embeddable* into (Y,d$'$) if there is a map f: X $\to$ Y such that d(x,y) = d$'$(f(x),f(y)) for all x and y in X.

As Example 4.1 illustrates, distortion is often necessary for embedding to occur. In that particular case, the distances can be distorted by a factor of $\sqrt{2}$ in order to form the square cycle.

Embedding a metric space in $\mathbb{R}^{\mathrm{N}}$ is a useful case of embedding, but embedding can be described in general settings.

Definition 4.5. For $0 < p < \infty$, ℓ_p *space* is the set of all real sequences $\{x_{\mathrm{n}}\}$ such that $\sum_{\mathrm{n}} |x_{\mathrm{n}}|^p < \infty$.

The norm of this space is given by

$$\|x\|_p = (\sum_{\mathrm{n}} |x_{\mathrm{n}}|^p)^{\frac{1}{p}}$$

Note that when p = 2 this is the Euclidean norm.

Definition 4.6. A metric space (X,d) is ℓ_p *embeddable* if (X,d) is isometrically embeddable into ℓ_p^n for some natural number n. This number n is the ℓ_p *dimension* of (X,d).

4.1. Embedding in ℓ_2.

Embedding in ℓ_2 attracts special attention. To those looking to analyze large amounts of data, translating data points into a finite metric space and then into a representation can be useful. In ℓ_2 there are extremely well developed tools in analysis and geometry to aid in the analysis of the data, so obtaining a good representation is important.

For its usefulness, ℓ_2 is very strict in its behavior, making embeddings difficult. The general theory of Banach spaces gives additional insight into why this is the case and additional motivation to consider ℓ_2 embeddings.

Definition 4.7. The *Banach−Mazur distance* is a measure of distance on the set of n-dimensional normed spaces. Take two normed spaces X and Y of dimension n and $\mathrm{GL}_{\mathrm{X,Y}}$, the set of linear isomorphisms from X to Y.

The *Banach−Mazur distance* between X and Y is defined to be

$$\delta(\mathrm{X,Y}) = \log(\inf_{\mathrm{T} \in GL_{\mathrm{X,Y}}} \mathrm{distortion(T)})$$

This is a metric on the space of n-dimensional normed spaces.

For many purposes (including ours) the *multiplicative Banach−Mazur distance*

$$\mathrm{d(X,Y)} = \mathrm{e}^{\delta(X,Y)} = \inf_{\mathrm{T} \in GL_{\mathrm{n}}} \mathrm{distortion(T)}$$

will be used. Because δ(X,Y) is a metric, the multiplicative Banach−Mazur distance obeys the multiplicative triangle inequality, d(X,Z) $\leq$ d(X,Y)d(Y,Z).

For convenience, this will be referred to as the *Banach−Mazur distance.*

The Banach−Mazur distance gives a sense of how close two normed spaces are to one another. If the distance is small, then the space needs little distortion for there to be a linear isomorphism between them. The following theorem, Dvoretzky's theorem, is a classical theorem which gives a quantitative sense of how close ℓ_2 space is to arbitrary normed spaces.

Theorem 4.8. *(Dvoretzky's Theorem* [10]*) For every $n \in \mathbb{N}$ and $\epsilon > 0$, every n-dimensional normed space contains a subspace X of dimension $m = \Omega(\epsilon^2 log(n))$ such that $d(X,\ell_2) \leq 1 + \epsilon$.*

Ω denotes that m is bounded asymptotically by $\epsilon^2 log(n)$ as $n \to \infty$.

4.1.1. *Bourgain's Theorem.* [15]

Motivated by this property of ℓ_2, in 1986, Jean Bourgain developed an algorithm which describes embedding in ℓ_2.

Theorem 4.9. *Any metric space (X,d) with n points can be embedded in ℓ_2 with distortion $\leq$ O(log n).*

Proof. Bourgain's proof gives an efficient randomized algorithm for the embedding in ℓ_2 with distortion $\leq$ O(log n).

Take a metric space (X,d) with n points.

1. Take m and q to be integers m = $\lfloor \log_2 \rfloor$ and q = $\lfloor \text{Clog}(n) \rfloor$ where C is a constant.
2. Construct an embedding into ℓ_2^{mq} with coordinates i = 1,...,m and j = 1,...,q.
3. Construct subsets of X, A_{ij} by putting each $x \in X$ into A_{ij} with probability 2^{-j}.
4. Now embed with function $f(x)_{ij} = d(x, A_{ij})$.

This is an embedding in $\ell_2^{O(\log)^2 n}$. It has distortion O(log n). $\square$

4.1.2. *Tightness of Bound.*

The construction of this algorithm raises the question whether a better embedding can be achieved. A paper by Nathan Linial (2002) shows that this bound is tight. He considers a specific type of graph that has a shortest path metric which is as far from the ℓ_2 metric as possible in order to guarantee a large distortion, giving a lower bound on distortion of graphs. To state his theorem, some definitions from graph theory are be needed.

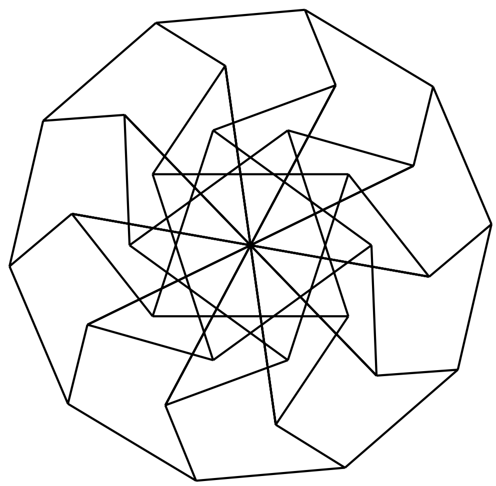
Definition 4.10. The *girth* of a graph is the shortest cycle contained in the graph. The girth of an acyclic graph is defined to be infinite.

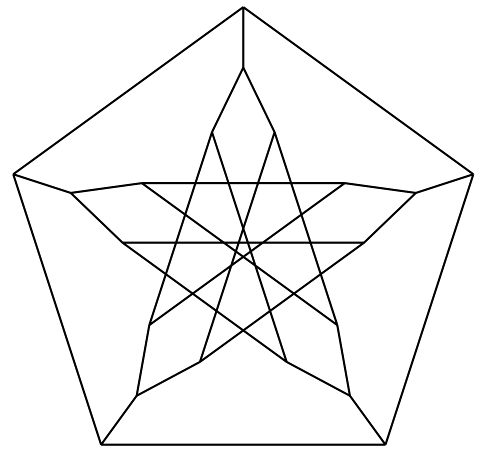
Definition 4.11. An *expander graph* is a connected graph in which every "small" subset of vertices has a "large" boundary. That is, the graph cannot be disconnected without removing many edges.

This quality can be quantified in the notion of an ϵ *edge expander*. A graph with n vertices is an ϵ edge expander if every set of K vertices with $0 \leq K \leq \frac{n}{2}$ has $\epsilon|K|$ edges connected to K^c (the set of vertices not in K).

Definition 4.12. A *k−regular graph* is a graph where each vertex is of degree k.

Example 4.13. Two instances of 3−regular graphs





Theorem 4.14. *Linial's Lower Bound* [10]

Take G, a k−regular graph, with $k \geq 3$, and girth g. Then every embedding $f : G \to \ell_2$ has distortion $\Omega(\sqrt{g})$.

Proof Sketch. This proof uses a random walk on the graph. Knowing the girth of the graph and that all vertices are connected to k other vertices, it can be proven that the walk moves away from where it started at constant speed at a time bounded asymptotically by g.

The geometry of Euclidean space means that this class of random walks is at time T expected to be $O(\sqrt{T})$ from its origin.

This difference must be accounted for by a distortion in the metric if it is to be embedded in ℓ_2. Comparing the two walks on the graph at time O(g) gives a distortion of $\Omega(\sqrt{g})$

This result is illustrated by the examples above. The triangle inequality is satisfied by equality many times, necessitating significant distortion.

4.1.3. *Isometric Embedding in ℓ_2.*

I. J. Schoenberg's 1937 paper [7] outlines the necessary and sufficient conditions for an isometric embedding in ℓ_2. He addresses separable pseudometric spaces. He characterizes embeddable metrics in terms of positive definite functions.

Definition 4.15. A real function $f = f(x_1, x_2, ..., x_n)$ is a *positive definite function* if it is defined for all real values and if for any real numbers $x_1, x_2, ..., x_n$ the $n \times n$ matrix A, where $A = (a_{i,j})$ and $a_{i,j} = f(x_i - x_j)$ is a positive, semi-definite matrix (that is, $x^t A x \geq 0$ for all real numbers x).

A similar notion of *positive definite functions* can be defined for real-valued functions which take as input distances on a pseudometric space (X,d).

A real function g(t) is *positive definite* if g is continuous, even, defined on the range of distances in the pseudometric space and satisfies the inequality

$$\sum_{\text{i,j} = 1}^{n} g(d(x_i, x_j)) \geq 0$$

An example of a positive definite function in ℓ_2 is function $f(t) = e^{-t^2}$ as is $e^{-\lambda t^2}$ for all real λ.

Theorem 4.16. *Schoenberg's Embedding*

Take separable pseudometric space (X,d). It is isometrically embeddable in ℓ_2 if and only if the functions $e^{-\lambda t^2}$ are positive definite in (X,d).

Proof Sketch The idea of this proof is to note that $e^{-\lambda t^2}$ for ($\lambda \in \mathbb{R}$) is a family of positive definite functions in ℓ_2. It is only necessary to consider $\lambda > 0$ as $\lambda = 0$ is an accumulation point of this family and the cases where $\lambda < 0$ follow by symmetry. The proof uses ideas from analysis about positive definite functions to show that if the given characteristics of positive definite functions are preserved on embedding into ℓ_2, then all distances must have been preserved and if the given family of functions are positive definite in the metric space, then the metric of the space will allow isometric embedding into ℓ_2.

4.2. Embedding in ℓ_1.

Following the formula given for ℓ_p space ℓ_1 is the set of all real sequences $\{x_n\}$ such that $\sum_n |x_n| < \infty$.

Define the distance on ℓ_1 as $d_{\ell_1}(x,y) = \sum_n |x_n - y_n| < \infty$.

To consider isometric embedding in ℓ_1, the *cut semimetric* will be used.

Definition 4.17. The *cut semimetric* is a pseudometric d on a set X. Given partitions A and B of X, define d(x,y) = 0 if x,y $\in$ A or x,y $\in$ B and define d(x,y) = 1 otherwise.

Every cut semimetric is clearly isometrically embeddable in ℓ_1.

The set of all linear combinations of semimetrics on a set forms a special class of metrics on that set. These are exactly the ℓ_1 metrics on the set (that is, the metrics which can be isometrically embedded in ℓ_1) [4].

4.3. Embedding in ℓ_∞.

Definition 4.18. ℓ_∞ space is defined to be the set of all real bounded sequences. It takes on the norm $\|x\|_\infty = \sup\limits_{n\in\mathbb{N}}|x_n|$.

Theorem 4.19. [1] *Every finite metric space (X,d) with n points can be embedded in ℓ_∞^n.*

Proof. Take a finite metric space (X,d) with X = $\{x_1,x_2,...,x_n\}$. Define an embedding function f: X $\to \ell_\infty^n$ by $f(x_i)_j = d(x_i,x_j)$ $\forall$ $1 \le i, j \le n$. ☐

Embeddings into lower dimensional ℓ_∞^k spaces exist.

Definition 4.20. Take a metric space (X,d) and every subset S $\subset$ X. Then define a mapping $f_S : X \to \mathbb{R}$ for each S by

$$f_S(x) = d(x, S) = \min_{s\in S} d(x, s)$$

A *Frechet Embedding* is a map $f : X \to \mathbb{R}^k$ where each coordinate in $\mathbb{R}^k$ is a scaled f_S mapping. f is then a *Frechet Embedding* if, for some $\beta_S \in \mathbb{R}$

$$f(x) = \bigoplus_{S\subset X} \beta_S f_S(x)$$

Proposition 4.21. [18] *When $\beta_S = 1$ $\forall S \subset V$, $||f(x) - f(y)||_\infty \le d(x, y)$. That is, Frechet embeddings are contraction mappings in the ℓ_∞ metric.*

Proof. Let S_x denote the point in $S \subset X$ closest to some point $x \in X$. Then both

$$d(x,S) - d(y,S) \leq d(x,S_y) - d(y,S_x) \leq d(x,y)$$

$$d(y,S) - d(x,S) \leq d(y,S_x) - d(x,S_y) \leq d(x,y)$$

This implies that

$$||f(x) - f(y)||_\infty = |d(x,S) - d(y,S)| \leq d(x,y)$$

□

A 1996 paper by Jiri Matousek uses these mappings to do distorted mappings into lower dimension ℓ_∞^k space.

Theorem 4.22. [23] *Take an n−point metric space (X,d) and integer D. Then (X,d) can be embedded into* $\ell_\infty^{O(Dn^{\frac{2}{D}} log(n))}$.

Proof Sketch The idea of this proof is to divide X into $O(Dn^{\frac{2}{D}} log(\mathrm{n}))$ subsets, each of which will correspond to a dimension in the range ℓ_∞ space.

Construct embedding function $\psi : (X,d) \to \ell_\infty^{O(Dn^{\frac{2}{D}} log(\mathrm{n}))}$ to be a Frechet embedding with j^{th} coordinate of $\psi(x)$ to be $d(x,S)$. Noting the proposition above, function ψ must be a contraction mapping. The rest of the proof uses an algorithm and probability to show that its contraction is limited.

4.4. Embedding in $\mathbb{R}^{\mathbf{N}}$. [8]

A paper by C.L. Morgan published in 1974 proved necessary and sufficient conditions for embedding a metric space in $\mathbb{R}^{\mathrm{N}}$. His theorem applies to arbitrary metric spaces, not only finite ones. It holds special interest for embedding finite metric spaces. His theorem makes the computation necessary to determine whether embeddability is feasible. His proof also shows that for any metric space, embedding into $\mathbb{R}^{\mathrm{N}}$ is a very strong condition, but it is one that is determined by a finite number of points in the metric space.

In order to state and prove the embedding theorem, some special definitions will be needed, as well as some general results about inner products, metrics, and linear algebra.

Definition 4.23. An *inner product* on a vector space V over a field F with characteristic 0 is a bilinear map $< \ , \ >: V \times V \to F$. This function satisfies conjugate symmetry and positive definiteness.

For a vector space V with element $x \in V$, define a norm $||x|| = \sqrt{< x, x >}$.

Theorem 4.24. *For a vector space V over characteristic 0 field F with inner product* $< \ , \ >$ *and norm* $||x|| = \sqrt{< x,x >}$, *a metric d(x,y)* $= ||x-y||$ *is induced by the norm.*

Definition 4.25. Take metric space (X,d) and points $x,y,z \in X$. Then define a function from $X \times X \times X \to \mathbb{R}$ as follows

$$< x, y, z >= \frac{1}{2}(d(x,z)^2 + d(y,z)^2 - d(x,y)^2)$$

If we set X to be a subset of some vector space V such that metric d is *induced* by an inner product on V, then $< x, y, z >$ would be the inner product of x−z and y−z.

Definition 4.26. Take metric space (X,d). Then set Y is a *metric subspace* of X if Y $\subset$ X and Y has the distance function d$|_{Y\times Y}$.

Finite metric subspaces of X are *$n-$simplices* in X. In particular, a metric subspace of n + 1 elements is an *$n-$simplex in X*.

If (X,d) is a subspace of Euclidean space, then these simplices have a clear notion of volume. The following function with begin to generalize this idea to arbitrary metric spaces.

Definition 4.27. Define a function D: $X^{n+1} \to \mathbb{R}$ as follows

Construct an n $\times$ n matrix, A, from (x_0,x_1,...,x_n) with real entries ($a_{i,j}$) = $< x_i, x_j, x_0 >$

Let D(x_0,x_1,...,x_n) = det(A).

This function D is a real valued function on the n-simplices of X.

Proposition 4.28. *The function D is symmetric.*

In Euclidean space, the entry ($a_{i,j}$) in the above matrix is

$$< x_i, x_j, x_0 >= \frac{1}{2}((\sqrt{(x_i - x_0)^2})^2 + (\sqrt{(x_j - x_0)^2})^2 - (\sqrt{(x_i - x_j)^2})^2)$$

$$= \frac{1}{2}((x_i - x_0)^2 + (x_j - x_0)^2 - (x_i - x_j)^2) = \frac{1}{2}(-2x_jx_0 - -2x_0x_i + 2x_ix_j + 2x_0^2)$$

$$\frac{1}{2}(-2x_jx_0 - 2x_0x_i + 2x_ix_j + 2x_0^2) = -x_jx_0 - x_0x_i + x_jx_i + x_0^2 = (x_i - x_0) * (x_j - x_0)$$

The determinant of a matrix with these entries is the square of the volume of a parallelpiped spanned by the set of n vectors (x_1,...,x_n) based at x_0.

With this machinery, it is possible to find the volume of the simplex (x_0,x_1,...,x_n).

Proposition 4.29. *The volume of the n-simplex Y = (x_0,x_1,...,x_n) in Eucledian space is*

$$Vol_n(Y) = \frac{1}{n!}\sqrt{D(x_0, x_1, ..., x_n)}$$

Having computed this volume in Euclidean space, define the volume of an n-simplex Y in any metric space to be the formula given by $\mathrm{Vol}_n(Y)$.

We can now provide two definitions which will describe which metric spaces can be embedded in $\mathbb{R}^N$

Definition 4.30. Metric space (X,d) is *flat* if each n-simplex Y in X, $\mathrm{Vol}_n(Y)$ is real.

Definition 4.31. Take flat metric space (X,d). The *dimension* of (X,d) is the largest $n \in \mathbb{N}$ such that there is an n-simplex of X with positive volume.

These characteristic of metric spaces will determine which can be isometrically embedded in $\mathbb{R}^N$. To prove Morgan's main theorem, some results from linear algebra are needed.

Lemma 4.32. *Any real n-dimensional inner product space is linearly isometric to Euclidean n-space.*

Lemma 4.33. *Let M be an m $\times$ m real symmetric matrix with all non-negative eigenvalues.*

Define $D[i,j]$ be the determinant of the $m-1 \times m-1$ minor of M obtained by deleting its ith row and jth column. Then $D[i,j]^2 \leq D[i,i]D[j,j]$

Theorem 4.34. *Morgan's Embedding in $\mathbb{R}^N$. A metric space can be isometrically embedded in Euclidean n-space iff the metric space is flat and has dimension less than or equal to n.*

Proof. Take a metric space (X,d) which can be isometrically embedded in Euclidean n-space. Isometries preserve volume, so the simplices must have real volume in (X,d) (as they have real volume in $\mathbb{R}^{\mathrm{N}}$), so (X,d) is flat. Because volume is preserved, the simplices of positive volume in (X,d) have positive volume in $\mathbb{R}^{\mathrm{N}}$, because there cannot be any simplices of positive volume in $\mathbb{R}^{\mathrm{N}}$ with greater than n + 1 points, (X,d) must have dimension less than or equal to n.

Take metric space (X,d) which is flat and of dimension n and n-simplex Y = (x_0,x_1,...,x_n) such that Y has positive volume.

If a map f: $X \to \mathbb{R}^{\mathrm{N}}$ can be constructed such that f embeds X isometrically in $\mathbb{R}^{\mathrm{N}}$ with some inner product, then because any real n-dimensional inner product space is linearly isometric to Euclidean n-space, (X, d) can be embedded in Euclidean n-space.

Define f: $X \to \mathbb{R}^{\mathrm{N}}$ as follows

$$f(x) = (< x, x_1, x_0 >, ..., < x, x_n, x_0 >)$$

Define a bilinear form on $\mathbb{R}^{\mathrm{N}}$ as follows. Take n × n matrix L with entries $(a_{i,j}) =< x_i, x_j, x_0 >$. Define bilinear form

$$< u, v >= u^t L^{-1} v, \ \forall u, v \in \mathbb{R}^{\mathrm{N}}$$

The claim is that this bilinear form is an inner product on $\mathbb{R}^{\mathrm{N}}$ and that f embeds (X,d) isometrically into this inner−product space. This is true if the eigenvalues of matrix L are positive.

Consider the polynomial det(xI + L). Its roots are the negatives of the eigenvalues of L. Look at the coefficient of the term of degree $n-k$ in this polynomial. It is the sum of the k*n minors of L which lie along the main diagonal. These minors are all non-negative because they are volumes of k-simplicial complexes (these volumes are all real, nonnegative as (X,d) is flat and dimension n). These make the polynomial positive, so it must have no positive roots, so there cannot be negative eigenvalues of L. L being symmetric and non-singular (as it (X,d) has non-zero dimension) ensures that its eigenvalues are positive.

To show that embedding function f is an isometry, it must be shown that f preserves the structure of (X,d). If this is true, then the inner product given on $\mathbb{R}^{\mathrm{N}}$ preserves the structure of all of the n-simplexes of (X,d). Thus it suffices to show that

$$< f(x), f(y) >=< x, y, x_0 >$$

for all x,y in X.

Construct a $(n+2) \times (n+2)$ matrix M with entries $< x_j, x_i, x_0 >$. By the same reasoning used on the similarly constructed matrix L, M has all non-negative eigenvalues.

Set $D[i,j]$ to be the determinant of the $(n+1) \times (n+1)$ of the matrix obtained by deleting the i^{th} row and j^{th} column of M.

Recall the lemma stating that

$$D[i,j]^2 \leq D[i,i]D[j,j]$$

$D[i,i]$ is the determinant corresponding to the volume of a $(n+1)-$simplex squared and scaled by a factor of $(n+1)!$. (X,d) is n-dimensional, so the volume of any $(n+1)-$simplex must be zero, so $D[i,i] = 0$. By the lemma, this means that $D[i,j] = 0$.

Setting $i = n$ and $j = n+1$ shows that, in particular, $D[n, n+1] = 0$.

Consider the minor of M with the nth row and (n + 1)st columns deleted.

$$\begin{bmatrix} <x_1,x_1,x_0> & \dots & \dots & <x_n,x_1,x_0> & <x_{n+2},x_1,x_0> \\ \vdots & \ddots & \ddots & \dots & \vdots \\ <x_1,x_{n-1},x_0> & \dots & \dots & <x_n,x_{n-1},x_0> & <x_{n+2},x_{n-1},x_0> \\ <x_1,x_{n+1},x_0> & \dots & \dots & <x_n,x_{n+1},x_0> & <x_{n+2},x_{n+1},x_0> \\ <x_1,x_{n+2},x_0> & \dots & \dots & <x_n,x_{n+2},x_0> & <x_{n+2},x_{n+2},x_0> \end{bmatrix}$$

Note that by the definition of the inner product

$$< f(x), f(y) >= f(x)^t L^{-1} f(y)$$

The condition for isometry is

$$< f(x), f(y) >=< x, y, x_0 >$$

Set x = x$_{n+1}$ and y = x$_{n+2}$ so that

$$f(x) = (< x_{n+1}, x_1, x_0 >, ..., < x_{n+1}, x_n, x_0 >)$$

$$f(y) = (< x_{n+2}, x_1, x_0 >, ..., < x_{n+2}, x_n, x_0 >)$$

Note that by deleting one row and one column from the matrix above, and dividing by the determinant of L, the matrix becomes the L^{-1} (when assigning the correct cofactor signs).

Expand the above matrix by the last row to calculate the determinant, using the minors

$$\begin{bmatrix} <x_1,x_1,x_0> & \dots & \dots & <x_n,x_1,x_0> \\ \vdots & \ddots & \ddots & \dots \\ <x_1,x_{n-1},x_0> & \dots & \dots & <x_n,x_{n-1},x_0> \\ <x_1,x_{n+1},x_0> & \dots & \dots & <x_n,x_{n+1},x_0> \end{bmatrix}$$

...

$$\begin{bmatrix} <x_2,x_1,x_0> & \dots & \dots & <x_{n+2},x_1,x_0> \\ \vdots & \ddots & \ddots & \dots \\ <x_2,x_{n-1},x_0> & \dots & \dots & <x_{n+2},x_{n-1},x_0> \\ <x_2,x_{n+1},x_0> & \dots & \dots & <x_{n+2},x_{n+1},x_0> \end{bmatrix}$$

Taking the appropriate sign changes and summing their determinants gives zero (as $D[n, n+1] = 0$). So dividing by det(L) still yields zero.

Continue the calculation to get that

$$< x_{n+1}, x_{n+2}, x_0 >= f(x_{n+1})^t L^{-1} f(x_{n+2})$$

This means that

$$< f(x), f(y) >=< x, y, x_0 >$$

for all x, y in X. This means that f is an isometry. □

These characterizations of metric spaces provides a useful way to analyze examples of metric spaces.

Theorem 4.35. [8] *For $n \geq 2$, $\mathbb{R}^N$ with the ℓ^p metric is flat iff $p = 2$.*

Proof. Morgan gives the two examples used below for his proof of this theorem without additional argument. However, working through the process to show why these examples work illustrates why the case when p = 2 is special.

Given $\mathbb{R}^{\mathrm{N}}$ with the ℓ^2 metric, the previous theorem proves that it is flat (i.e. ($\mathbb{R}^{\mathrm{N}}$,ℓ^2) can embed in itself).

The example given in 4.1 of a non-embeddable metric space suggests how to construct simplices of imaginary volume in ($\mathbb{R}^{\mathrm{N}}$,ℓ^p) when p $\neq$ 2. It is only necessary to find examples in $\mathbb{R}^2$ as $\mathbb{R}^2 \subset \mathbb{R}^{\mathrm{N}}$ for n $\geq$ 2.

Consider ($\mathbb{R}^{\mathrm{N}}$,ℓ^p) for p < 2.

If $1 \leq p$ the ℓ^p metric is induced by the norm

$$\|x\|_p = (\sum_{\mathrm{n}} |x_{\mathrm{n}}|^p)^{\frac{1}{p}}$$

Take example of the 3-simplex Y in ($\mathbb{R}^{\mathrm{N}}$,ℓ^p) with Y = {(0,0),(1,0),(1,1),(0,1)}.

Observe that for any value of p $\geq$ 1, the horizontal and vertical distances on this simplex are the same.

If p $\geq$ 1,

$$d((a,b),(a,c)) = \|(a,b)-(a,c)\|_p = (|(a-a)|^p + |(b-c)|^p)^{\frac{1}{p}} = |b-c|$$

The same argument applies, by symmetry, when the second coordinates are equal.

This means that distortion would occur in the distance between two non-adjacent points in this simplex.

By the triangle inequality, for any p $\geq$ 1

$$d((0,0),(1,1)) \leq d((0,0),(0,1)) + d((0,1),(1,1)) = 1+1 = 2$$
$$d((0,0),(1,1)) \leq d((0,0),(1,0)) + d((1,0),(1,1)) = 1+1 = 2$$
$$d((0,0),(1,1)) = \|(0,0)-(1,1)\|_p = (|(0-1|^p + |(0-1)|^p)^{\frac{1}{p}} = 2^{\frac{1}{p}}$$

As $p \to \infty$, the quantity $d((0,0),(1,1)) \to 1$, so this square in ($\mathbb{R}^{\mathrm{N}}$,ℓ^2) collapses to a line as p increases.

Now consider the matrix constructed to compute function D(Y)

$$A = \begin{bmatrix} <(0,0),(1,0),(1,0)> & <(0,0),(1,0),(1,1)> & <(0,0),(1,0),(0,1)> \\ <(0,0),(1,1),(1,0)> & <(0,0),(1,1),(1,1)> & <(0,0),(1,1),(0,1)> \\ <(0,0),(0,1),(1,0)> & <(0,0),(0,1),(1,1)> & <(0,0),(0,1),(0,1)> \end{bmatrix}$$

Using the formula

$$<x,y,z> = \frac{1}{2}(d(x,z)^2 + d(y,z)^2 - d(x,y)^2)$$

Any entry on the diagonal takes the form

$$<x,y,y> = \frac{1}{2}(d(x,y)^2 + d(y,y)^2 - d(x,y)^2) = 0$$

A has a zero diagonal.

$$A = \begin{bmatrix} 0 & <(0,0),(1,0),(1,1)> & <(0,0),(1,0),(0,1)> \\ <(0,0),(1,1),(1,0)> & 0 & <(0,0),(1,1),(0,1)> \\ <(0,0),(0,1),(1,0)> & <(0,0),(0,1),(1,1)> & 0 \end{bmatrix}$$

Using the fact that for any p value,

$$d((0,0),(0,1)) = d((0,0),(1,0)) = d((1,0),(1,1)) = d((0,1),(1,1)) = 1$$

Matrix A can be simplified to

$$A = \begin{bmatrix} 0 & \frac{1}{2}d((0,0),(1,1))^2 & \frac{1}{2}d((1,0),(0,1))^2 \\ 1-\frac{1}{2}d((0,0),(1,1))^2 & 0 & 1-\frac{1}{2}d((0,0),(1,1))^2 \\ \frac{1}{2}d((0,1),(1,0))^2 & \frac{1}{2}d((0,0),(1,1))^2 & 0 \end{bmatrix}$$

Then D((0,0),(1,0),(1,1),(0,1)) can be calculated

$$D(Y) = (\frac{1}{2}d((0,0),(1,1))^2)(\frac{1}{2}d((0,1),(1,0))^2)(1-\frac{1}{2}d((0,0),(1,1))^2)+$$

$$(\frac{1}{2}d((1,0),(0,1))^2)(1-\frac{1}{2}d((0,0),(1,1))^2)(\frac{1}{2}d((0,0),(1,1))^2)$$

This simplifies to

$$d((1,0),(0,1))^2 d((0,0),(1,1))^2(\frac{1}{2}-\frac{1}{4}d((0,0),(1,1))^2)$$

The term $d((1,0),(0,1))^2 d((0,0),(1,1))^2$ is always positive. This value of D(Y) is negative (and so the volume of Y imaginary) only when

$$\frac{1}{2} < \frac{1}{4}d((0,0),(1,1))^2$$

Solving this inequalities gives that the volume is imaginary when

$$\sqrt{2} < d((0,0),(1,1))$$

If $0 < p < 1$, then ℓ^p has metric

$$d_p(x,y) = \sum_{\mathrm{i}\,=\,1}^{n} |x_i - y_i|^p$$

Then

$$d((0,0),(1,1)) = \sum_{\mathrm{i}\,=\,1}^{2} |0-1|^p = 1^p + 1^p = 2$$

D(Y) is negative for $0 < p < 1$, so Vol(Y) is imaginary, so $(\mathbb{R}^{\mathrm{N}}, \ell^p)$ is not flat for $0 < p < 1$. If $1 \leq p < 2$, then this distance takes the form

$$d((0,0),(1,1)) = \|(0,0)-(1,1)\|_p = (1^p + 1^p)^{\frac{1}{p}} = 2^{\frac{1}{p}}$$

If $p < 2$, then the inequality is satisfied, meaning that $(\mathbb{R}^{\mathrm{N}}, \ell^p)$ is not flat for $1 \leq p < 2$.

Consider $(\mathbb{R}^{\mathrm{N}},\ell^p)$ for p > 2. Take example of the 3-simplex Y in $(\mathbb{R}^{\mathrm{N}},\ell^p)$ with Y = {(0,1),(1,0),(-1,0),(0,-1)}. This simplex has vertical and horizontal distances of 2 which are preserved in all $(\mathbb{R}^{\mathrm{N}},\ell^p)$ for all p. It is the distances which are not preserved which will cause this simplex to have imaginary volume for $p > 2$. This example's invariant distances are larger than the changing distances, so by repeating the same computation as above, the inequality is reversed, giving that the volume of Y is imaginary when

$$\sqrt{2} > d((-1,0),(0,1))$$

This is an equality when p = 2. By the same analysis as above, as p becomes greater than 2, this inequality is satisfied, showing that Y has an imaginary volume when $p > 2$. This means that $(\mathbb{R}^{\mathrm{N}},\ell^p)$ is not flat for $p > 2$. □

4.5. **Embeddings of the ℓ_2 Metric.**

In section 4.2 it was shown that ℓ^2 is close to other normed spaces, that is, there is a linear isomorphism between them which requires little distortion of the spaces. It is then natural to ask when there is an isometric embedding from ℓ_2 to other spaces.

4.5.1. *Dimension Reduction in ℓ_2.*

Given a metric space (X,ℓ_2) in $\mathbb{R}^{\mathrm{N}}$, it is useful to ask whether the dimension of the host space, ℓ_2, can be reduced in exchange for distortion. A paper by William Johnson and Joram Lindenstrauss quantified the possible dimension reduction.

Theorem 4.36. *(Johnson and Lindenstrauss Dimension Reduction* [19]*) Given any n point metric space (X,ℓ_2) $\subset \mathbb{R}^N$ and $\epsilon > 0$, there is an embedding of distortion of at most $1+\epsilon$ such that*

$$(X, \ell_2) \to \ell_2^{O(\frac{log\ n}{\epsilon^2})}$$

The proof of this dimension reduction theorem and other proofs of isometric embedding from ℓ_2 to ℓ_p uses a technique in theoretical computer science, *random projection.*

Definition 4.37. Take vectors r_1,...,$\mathrm{r}_k \subset \mathbb{R}^{\mathrm{N}}$ which have been obtained by some random process. Then define map $\psi : \mathbb{R}^N \to \mathbb{R}^k$ as follows

$$\psi : v \to (< v, r_1 >, ..., < v, r_k >)$$

Map ψ is a *random projection* from $\mathbb{R}^{\mathrm{N}} \to \mathbb{R}^k$.

Random projection ψ can be conveniently expressed as a $k \times n$ matrix A whose rows are r_1,...,r_k so that $\psi(v) = Av$. This means that random projections are linear.

There are three notable examples of random process used to generate the r_1,...,r_k. All three have been used to prove the Johnson$-$Lindenstrauss Theorem.

Examples 4.38. .

1. Set each $\mathrm{r}_i = (\mathrm{r}_i^1,...,\mathrm{r}_i^n)$. Obtain values for each r_i^j from a normal probability distribution centered at 0 with variance 1. This is labeled ψ_N and was used to prove Johnson$-$Lindenstrauss [21].
2. Set each $\mathrm{r}_i = (\mathrm{r}_i^1,...,\mathrm{r}_i^n)$. Obtain values for each r_i^j by choosing either +1 or -1, each with probability $\frac{1}{2}$. This method is called *binary coins.* This is labeled ψ_B. This is the simplest method used to prove Johnson$-$Lindenstrauss [22].
3. Take r_1,...,r_k to be a set of k orthogonal vectors from $\mathbb{S}^{n-1}$. This is labeled ψ_S and was originally used by Johnson and Lindenstrauss [19].

4.5.2. *Isometric Embedding from ℓ_2 to ℓ_1.*

Two interesting cases of ℓ_p spaces are ℓ_2 and ℓ_1, so the existence of an isometric embedding of a $n-$point metric space in ℓ_2^n to some finite dimensional ℓ_1^k is an important one. In order to prove that there does exist such an embedding, the space $\ell_1^{\mathbb{S}^{n-1}}$ will be explored. The definition of this space and the proof of an embedding theorem is given in lecture 12 of the series on finite metric spaces given at TTIC [18].

Definition 4.39. Space $\ell_1^{\mathbb{S}^{n-1}}$ is an ℓ_1 metric space with a coordinate for each vector in $\mathbb{S}^{n-1}$. Each point in $\ell_1^{\mathbb{S}^{n-1}}$ in given by a function f: $\mathbb{S}^{n-1} \to \mathbb{R}$. The ℓ_1 norm is then given by

$$||f||_1 = \int_{r \in \mathbb{S}^{n-1}} |f(r)| dr$$

Lemma 4.40. *There exists an isometric embedding of every $n-$point metric space in ℓ_2^n to $\ell_1^{\mathbb{S}^{n-1}}$.*

With this embedding lemma, it only need be shown that there is an isometric embedding from $\ell_1^{\mathbb{S}^{n-1}}$ into a finite dimensional ℓ_1. This result can also be generalized to isometric embeddings from $\ell_p^{\mathbb{S}^{n-1}}$ to finite dimensional ℓ_p

Theorem 4.41. [18] *Every $n-$point metric space in ℓ_2^n can be isometrically embedded into $\ell_1^{n!}$.*

Proof Sketch Isometrically embed space metric space X= $\{x_1,...,x_n\}$in ℓ_2^n by the above lemma. $\mathbb{S}^{n-1}$ is partitioned into n! regions and each region is assigned an x_i and x_j. Each region is defined in such a way that the sign of $< x_i, r > - < x_j, r >$ is constant within it. It can then be shown that this produces an isometric embedding from ℓ_2^n to $\ell_1^{\mathbb{S}^{n-1}}$ and into $\ell_1^{n!}$.

Acknowledgments. I would like to thank Peter May for suggesting this topic, giving enlightening lectures, and organizing the REU. I would also like to thank my mentors, Tori Akin and Lei Chen, for helping me put this paper together.

References

[1] Jacob Goodman. Handbook of Discrete and Computational Geometry. Chapter 8: Low Distortion Embeddings of Finite Metric Spaces. Piotr Indyk and Jiri Matousek. 2004.
[2] J. P. May. A Concise Course in Algebraic Topology. University of Chicago Press. 1999.
[3] J. P. May. Finite Spaces
[4] Michel Marie Deza. Geometry of Cuts and Metrics. 1-6,23-30,37-40. Springer-Verlag Berlin Heidelberg. 1997.
[5] Finite Metric Spaces & Their Embeddings: Introduction and Basic Tools. Manor Mendel. 2015
[6] Matrix Completions, Moments, and Sums of Hermitian Squares. Mihly Bakonyi & Hugo J. Woerdeman. Princeton University Press. 2011.
[7] Metric Spaces and Positive Definite Functions. I. J. Schoenberg. American Mathematical Society. 1937.
[8] Embedding Metric Spaces in Euclidean Space. C.L. Morgan. Journal of Geometry. 1974.
[9] Volume of an N-simplex by Multiple Integration. Richard S. Ellis. Birkhauser Mathematics. 1976.
[10] Finite Metric Spaces Combinatorics, Geometry and Algorithms. Nathan Linial. International Congress of Mathematicians. 2002.
[11] Lectures on Differential Geometry. Shlomo Sternberg. Chelsea Publishing Company. 1964.
[12] Principles of Mathematical Analysis. Walter Rudin. McGraw-Hill Education. 1976.
[13] Topology. James Munkres. Pearson. 2000.
[14] Uniform Embeddings of Metric Spaces and of Banach Spaces into Hilbert Spaces. I. Aharoni, B. Maurey, B.S. Mityagin. Israel Journal of Math. 1985.
[15] On Hilbertian Subsets of Finite Metric Spaces. J. Bourgain, T. Figiel, V. Milman. Israel Journal of Math. 1986.
[16] Euclidean Quotients of Finite Metric Spaces. Manor Mendel, Assaf Naor. Advances in Mathematics. 2004.
[17] Lower Bounds on the Distortion of Embedding Finite Metric Spaces in Graphs. Y. Rabinovich, R. Raz. Discrete & Computational Geometry. 1998

[18] Finite Metric Spaces. Lectures notes from Toyota Technological Institute at Chicago (TTIC) by Harald Rcke.

[19] Extensions of the Lipschitz Maps into a Hilbert space. William Johnson, Joram Lindenstrauss. Contemporary Mathematics. 1984.

[20] Embedding the Diamond Braph in L_p and Dimension Reduction in ℓ_1. James Lee and Assaf Naor. Geometric and Functional Analysis. 2004.

[21] Approximate Nearest Neighbors: Towards Removing the Curse of Dimensionality. Piotr Indyk and Rajeev Motwani. Proceedings of the 30th ACM Symposium on Theory of Computing. 1998.

[22] Database-Friendly Random Projections: Johnson-Lindenstrauss with Binary Coins. Dimitris Achlioptas. Journal of Computer and System Sciences. 2003.

[23] On the Distortion Required for Embedding Finite Metric Spaces into Normed Spaces. Jiri Matousek. Israel Journal of Mathematics. 1996.