University of New Hampshire
University of New Hampshire Scholars' Repository

Doctoral Dissertations

Student Scholarship

Spring 2019

Pseudometrics, The Complex of Ultrametrics, and Iterated Cycle Structures

Eric Robert Kehoe
University of New Hampshire, Durham

Pseudometrics, The Complex of Ultrametrics, and Iterated Cycle Structures

BY

Eric R. Kehoe

DISSERTATION

Submitted to the University of New Hampshire
in Partial Fulfillment of
the Requirements for the Degree of

Doctor of Philosophy
in
Mathematics

May, 2019

This thesis/dissertation has been examined and approved in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics by:

Thesis/Dissertation Director, David Feldman, Professor of Mathematics and Statistics

John Gibson, Professor of Mathematics and Statistics

Don Hadwin, Professor of Mathematics and Statistics

Maria Basterra, Professor of Mathematics and Statistics

Dmitri Nikshych, Professor of Mathematics and Statistics

On April 17th, 2019

Original approval signatures are on file with the University of New Hampshire Graduate School.

DEDICATION

To my love Emily.

ACKNOWLEDGEMENTS

I would like to thank my parents for always supporting my dream of becoming a mathematician. I would like to thank my family and friends for their support and reminding me to have fun. I would like to thank my band for giving me a creative outlet outside of academics.

I would like to thank the University of New Hampshire Mathematics and Statistics department and the Graduate School for providing me the resources to reach my potential as a mathematician. I would like to thank my dissertation advisor Professor David Feldman for teaching me how to ask the right questions, providing insight into mathematical writing, and most of all being a good friend.

I would like to thank my love Emily for being by my side throughout the writing of this dissertation.

PREFACE

Take a set X and the associated set of pseudometrics over X (here "pseudo" means the allowance of zero distances). If we impose the condition that the distances must be bounded by 1, we obtain the set of bounded-by-1 pseudometrics. At first glance, it may seem that the only extreme bounded-by-1 metrics on X are those given by partitions. That is metrics that only take the value 0 and 1. With the aid of computer experimentation and linear programming we see immediately that bounded-by-1 extreme metrics can have a wide variety of rational distances. Naturally we would like to classify such extreme metrics.

Before beginning the classification process, we try to learn what has been done for just pseudometrics. If X has cardinality n, it's set of pseudometrics form a convex cone in $\binom{n}{2}$-dimensional Euclidean space called the metric cone. There is a wealth of literature on the metric cone. In chapter 2 we give a brief account of the most important and relevant literature. Most notable, Avis classified many graphical extreme rays of the metric cone and showed that their local structure can be rationally arbitrary. Bendelt and Dress, whose motivation stems from the field of phylogeny, gave a canonical decomposition of metrics into bi-partition metrics (or splits). Using this decomposition evolutionary biologists are able to construct phylogenetic trees from dissimilarity data given by differences in the morphology of species. Thus far extreme rays of the cone have only been completely classified up to 6 points [14].

Characterizing extremeness for large classes of extreme metrics thus becomes desirable. Following Avis we give a new class of extreme rays, called bow-tie metrics, which exhibit a wide range of rational distances. In chapter 3 we begin our work with bounded-by-1 pseudometrics on X. Unlike the cone, bounded-by-1 pseudometrics over X form a convex **polytope** called the metric body. Extreme rays in the cone induce extreme points in the body, allowing us to transfer extreme data. Using the bow-tie metrics from chapter 2 we show that any separable bounded-by-1 pseudometric space can be extended to an extreme seperable bounded-by-1 pseudometric space. Hence, the local structure of a separable bounded-by-1 pseudometric space can be arbitrary (even with the inclusion of irrational distances).

The first extreme bounded-by-one pseudometric we encounter outside of the set of partitions is the so-called midpoint metric on 4 points. The midpoint metric is a metric which expresses a single point as the mid-point of triangle with edges of length 1. The midpoint metric has only distances 1 and $1/2$. Metrics with only these distances are called half-one metrics. In proving that the midpoint metric is extreme we discover a technique of proof.

In order to characterize the extremeness of a half-one metric we can build a graph, named the edge graph, whose edges are associated to the degenerate triangle inequalities for the metric. A half-one metric will then be extreme in the body if and only the components of its edge graph each contain at least one odd cycle.

Half-one metrics have received some attention. With probability limiting to 1 as n goes to infinity, half-ones optimize linear objective functions over $\mathbb{R}^{\binom{n}{2}}$. This raises the question of whether or not half-one metrics outnumber all other extreme bounded-by-1 pseudometrics. Building off the construction of bow-tie metrics and taking co-products one can produce large families of extreme metrics which rival the class half-one metrics in size.

A geometric explanation of their tendency to optimize follows. The probability that a corner on a convex polytope will optimize a linear objective function is proportional to the size the exterior solid angle at that corner. The solid exterior angle at a corner can be calculated as the volume of the dual cone at that corner intersected with the unit sphere centered at the corner. We conjecture that half-one metrics on the average have relatively large dual cones. For those familiar with the Gauss-Bonnet theorem, the convex polytope version of theorem replaces curvature on a manifold with the exterior solid polyhedral angles at vertices. So, from the optic of Gauss-Bonnet theorem, we mean to guess that extreme half-ones tend to optimize linear objection functions because they eat up almost all the curvature of the metric body.

In chapter 4 we treat half-one metrics as the primary object of interest. Every half-one metric induces an undirected **unweighted** graph given by its half-length edges. Using this graph we give a lower bound to the number of half-one metrics and show that half-one metrics outnumber partition metrics. Every metric with distances greater than or equal to a half is automatically a metric. These metrics form the upper half of the metric body. The upper half can be decomposed via decomposing non-extreme half-one metrics. The language of rigidity and perturbations provides a useful tool in the decomposition and we apply these concepts to find neighbors of extreme half-one metrics sitting on the boundary of the metric body.

In chapter 5 we give experimental results on the statistics of bounded-by-1 rational symmetric functions with a fixed denominator q, called q-level points. The goal of the experiments was to find large classes of q-level points in $\mathbb{R}^{\binom{n}{2}}$ that are extreme metrics with probability limiting to 1 as n goes to infinity. As a basis we start with 3-level points. This investigation led to revisiting the bowtie metric on 6 points. The bowtie generates a highly symmetric graphical extreme ray of the cone with denominator 3. By understanding completely why the bowtie metric is extreme, we are able to generalize its construction.

In chapter 6 we shift focus to a subclass of bounded-by-1 pseudometrics, the bounded-by-1 pseudoultrametrics. Ultrametrics satisfy a stronger version of the triangle inequality. Every tree metric generates an example of a bounded-by-1 pseudoultrametric. Bounded-by-1 pseudoultrametrics live in the convex hulls of partition chains. This fact enables us to determine the topology of so called bounded-by-1 pseudoultrametrics up to homotopy equivalence. Determining the homotopy type naturally leads to an investigation into the topology of the order complex of the partition lattice. With the help of SAGE, a mathematics software system, we found that the homology of the order complex was concentrated in the top dimension. This led to the idea that the complex was shellable, a sufficient property for a simplicial complex to be homotopy equivalent to a wedge of spheres.

It is a known fact that the order complex of the partition lattice is shellable. Indeed, Bjouner proved that any lattice that admits an L-labelling has a shellable order complex. To find the number of spheres in the decomposition one can calculate the Euler-characteristic of the order complex. Traditionally this is does by computing a certain value of the Möbius associated to the lattice. We improve upon the proof of the Euler-characteristic of the complex of scaled ultrametrics by giving a computable bijection on the faces of the complex. We then extend this technique of proof to the context of iterated cycle structures (ICS). ICS are new mathematical objects which naturally generalize partition chains in the lattice of partitions. Just as Stirling numbers of the second kind are dual to Stirling numbers of the first kind, the set of ICS are dual to the lattice of partitions. Future work involves finding a geometric realization of the "complex" of ICS.

Eric R. Kehoe

TABLE OF CONTENTS

ABSTRACT

Pseudometrics, The Complex of Ultrametrics,

and Iterated Cycle Structures

by

Eric R. Kehoe

University of New Hampshire, May, 2019

Every set X, finite of cardinality n say, carries a set $\mathcal{M}(X)$ of all possible pseudometrics. It is well known that $\mathcal{M}(X)$ forms a convex polyhedral cone whose faces correspond to triangle inequalities. Every point in a convex cone can be expressed as a conical sum of its extreme rays, hence the interest around discovering and classifying such rays. We shall give examples of extreme rays for $\mathcal{M}(X)$ exhibiting all integral edge lengths up to half the cardinality of X.

By intersecting the cone with the unit cube we obtain the convex polytope of bounded-by-one pseudometrics $\bar{\mathcal{M}}(X)$. Analogous to extreme rays, every point in a convex polytope arises as a convex combination of extreme points. Extreme rays of $\mathcal{M}(X)$ give rise to very special extreme points of $\bar{\mathcal{M}}(X)$ as we may normalize a nonzero pseudometric to make its largest distance 1. We shall give a simple and complete characterization of extremeness for metrics with only edge lengths equal to 1/2 and 1. Then we shall use this characterization to give a decomposition result for the upper half of the $\bar{\mathcal{M}}(X)$.

$\bar{\mathcal{M}}(X)$ contains the set of bounded-by-1 pseudoultrametrics, $U(X)$. Ultrametrics satisfy

a stronger version of the triangle inequality, and have an interesting structure expressed in terms of partition chains. We will describe the topology of U(X) and its subset of scaled ultrametrics, $\tilde{U}(X)$, up to homotopy equivalence. Every permutation on a set X can be written as a product of disjoint cycles that cover X. In this way, a permutation generalizes a partition. An iterated cycle structure (ICS) will then be the associated generalization of a partition chain. Analogously, we will compute the "Euler-characteristic" of the set of iterated cycle structures.

CHAPTER 1

Convex Geometry

This chapter develops the basic tools and language of convex geometry for application to the geometry of the set of metric spaces over a given set. Those familiar with these basic definitions and constructions may proceed directly to chapter 2.

1.1 Affine Spaces

Definition 1.1.1. An *affine subspace* V of $\mathbb{R}^m$ means the locus of points satisfying a linear (and generally inhomogeneous) equation $Ax = b$, A a real $r \times m$ matrix and $b \in \mathbb{R}^r$.

Definition 1.1.2. A set of points $X = \{x_1, \dots, x_n\} \subset \mathbb{R}^m$ is *affinely dependent* if there exist real numbers $\lambda_1, \dots, \lambda_n$ not all zero, such that

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = 0, \quad \lambda_1 + \cdots + \lambda_n = 0$$

and X is affinely independent if no such numbers exist.

Definition 1.1.3. An *affine combination* of points $X = \{x_1, \dots, x_n\} \subset \mathbb{R}^m$ is a linear combination $\lambda_1 x_1 + \cdots + \lambda_n x_n$ such that $\lambda_1 + \cdots + \lambda_n = 1$.

A set $S \in \mathbb{R}^m$ has *dimension* r, denoted $\dim(S) = r$, if a maximal affinely independent subset of S contains exactly $r + 1$ points.

Definition 1.1.4. The unique affine subspace $\mathrm{aff}(S)$ of smallest dimension containing S is called the *affine span* of S; $\mathrm{aff}(S)$ equals the set of all affine combinations of points in S. (Note that $\dim(S) := \dim(\mathrm{aff}(S))$.)

Within $\mathbb{R}^m$, we call affine subspaces with dimension $m-1$ (resp. 1 and 0), *hyperplanes* (resp. *lines* and *points*).

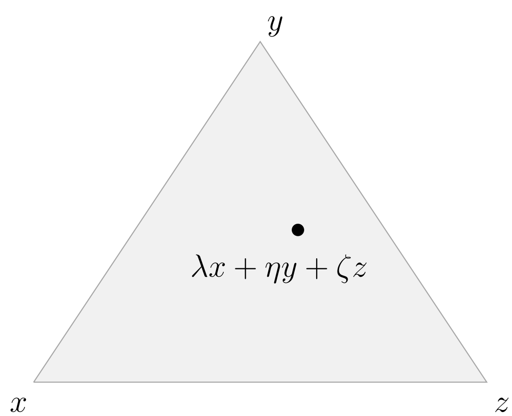
1.2 Convex Sets

Definition 1.2.1. A set $S \subset \mathbb{R}^m$ is *convex* if, given points $x, y \in S$, the interval

$$[x, y] := \{\lambda x + (1-\lambda)y \mid 0 \leq \lambda \leq 1\} \subseteq S.$$

Definition 1.2.2. A *convex combination* of points $X = \{x_1, \ldots, x_n\} \subset \mathbb{R}^m$ is an **affine** combination $\lambda_1 x_1 + \cdots \lambda_n x_n$ with all $\lambda_i \geq 0$.

Example 1.2.3. Convex Set in the $\mathbb{R}^2$



$$\lambda + \eta + \zeta = 1$$
$$\lambda, \eta, \zeta \geq 0$$

Definition 1.2.4. For any $S \subseteq \mathbb{R}^m$, the *convex hull*, $\mathrm{conv}(S)$, means the intersection of all convex sets containing S; $\mathrm{conv}(S)$ contains precisely all convex combinations of points from S. (Note that $\dim(\mathrm{conv}(S)) = \dim(S)$.)

Definition 1.2.5. A *convex polytope* P means the convex hull of a finite set of points $\{v_1, \ldots, v_n\} \subset \mathbb{R}^m$, so

$$P = \{\lambda_1 v_1 + \lambda_2 v_2 + \cdots \lambda_n v_n \mid \quad \forall i \lambda_i \geq 0, \quad \lambda_1 + \cdots + \lambda_n = 1\}$$

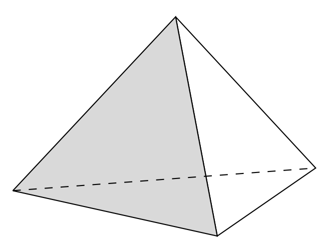
We call a polytope of dimension d a *d-polytope*.

We will speak of the *vertex representation* or V-rep of P when we give P as the convex hull of a minimal set of points. Non-trivially, every convex polytope equals some bounded intersection of finitely many closed half-spaces

$$P = \{x \in \mathbb{R}^m \,|\quad H \cdot x \leq b \quad \text{for some } H \in \mathbb{M}_{n,m}(\mathbb{R}) \text{ and } b \in \mathbb{R}^n\}\,;$$

we will speak of the *half-space representation* or H-rep when we've used a minimal set of half-spaces.

Example 1.2.6. Convex 3-Polytope in $\mathbb{R}^3$



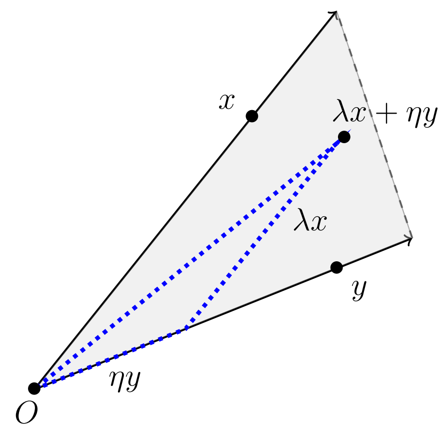
1.3 Conical Sets

Definition 1.3.1. A *conical combination* of points $X = \{x_1, \ldots, x_n\} \subset \mathbb{R}^m$ means a linear combination $\lambda_1 x_1 + \cdots \lambda_n x_n$ such that $\lambda_i \geq 0$.

Definition 1.3.2. Call a set $S \subset \mathbb{R}^m$ *conical* if, for any pair of points $x, y \in S$, the fan

$$R_{x,y} := \{\lambda x + \eta y \,|\, \lambda, \eta \geq 0\} \subseteq S.$$

Example 1.3.3. Conical Set in $\mathbb{R}^2$



Observe that conical implies convex.

Definition 1.3.4. For any set S, define the *conical hull* of S, $\text{cone}(S)$, as the intersection of all conical sets containing S; $\text{cone}(S)$ contains precisely all conical combinations of points from S. (Note that $\dim(\text{cone}(S)) = \dim(S \cup \{0\})$ equals the dimension of the smallest vector space containing S.)

Definition 1.3.5. A *pointed convex polyhedral cone* C means the conical hull of a finite set of points $\{v_1, \ldots, v_n\} \subset \mathbb{R}^m$ that lie in some closed half-space bounded by some hyperplane H through 0, but with not all the v_i in H itself. So

$$C = \{\lambda_1 v_1 + \lambda_2 v_2 + \cdots \lambda_n v_n \mid \quad \lambda_i \geq 0 \, \forall i\} \, .$$

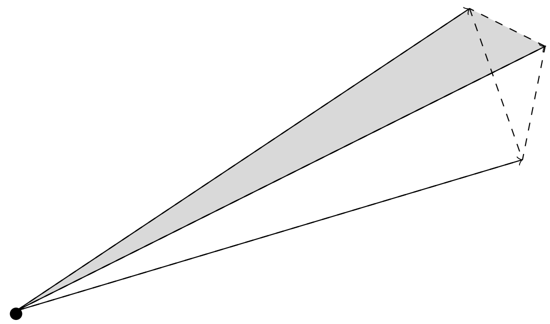
Call a polyhedral cone of dimension m an m-cone.

We speak of the *ray representation* or R-rep of C when we use a minimal set $\{v_1, \ldots, v_n\}$. Every pointed convex polyhedral cone also arises as the intersection of finitely many closed half-spaces with boundaries all passing through the origin, so

$$C = \{x \in \mathbb{R}^m \mid \quad H \cdot x \leq 0 \quad \text{for some } H \in \mathbb{M}_{n,m}(\mathbb{R})\}$$

and just as with polytopes, we speak of the half-space representation or H-rep when we've used a minimal set of half-spaces.

Example 1.3.6. Convex Polyhedral 3-Cone in $\mathbb{R}^3$



Definition 1.3.7. For a polyhedral cone C, the *dual cone* $C^\vee$ means the set

$$C^\vee = \{x \in \mathbb{R}^m \,|\, \forall y \in C \quad \langle x, y\rangle \geq 0\,\}$$

with $\langle \cdot, \cdot \rangle$ the standard inner product on $\mathbb{R}^m$.

We say a hyperplane H *supports* a closed convex set S if $H \cap S \neq \emptyset$ and S lies in exactly one of the two closed half-spaces bounded by H. If a hyperplane H supports S then we call $H \cap S$ a *face* of S. Every point in the boundary ∂S of S belongs to some supporting hyperplane of S, and thus lies in some face of S.

Theorem 1.3.8. *For an m-polytope (m-cone) S:*

1. *The faces of S constitute polytopes (polyhedral cones).*

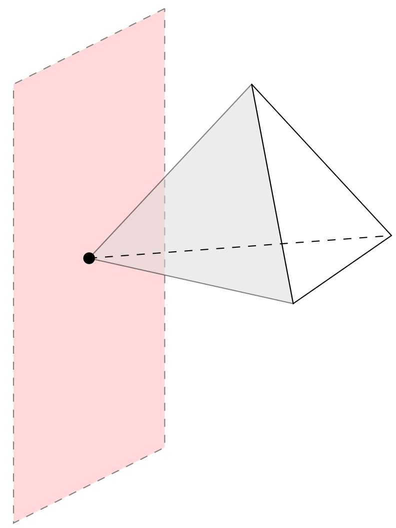
2. *S has faces of every dimension* $1 \leq d \leq m-1$.

3. *A face of a face of S is a face of S.*

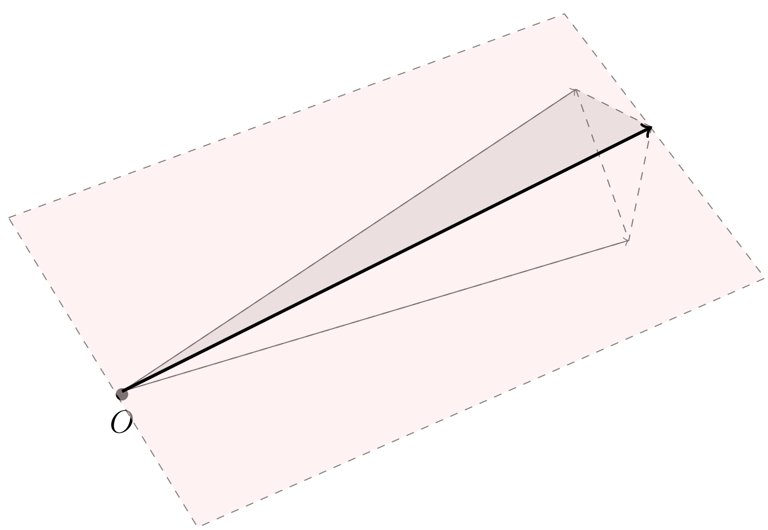
4. *S equals the convex (conical) hull of its extreme points (extreme rays)*

We call a 0-face a *vertex* or an *extreme point*, a 1-face an edge (*extreme ray*), and a $m-1$ face a *facet*. A cone S has only one vertex 0.

Example 1.3.9. Extreme point of a convex polytope



Example 1.3.10. Extreme ray of a convex polyhedral cone



Proposition 1.3.11. *For $S \subset \mathbb{R}^m$ a polytope, $z \in (x, y) \subset S$, and hyperplane H supporting S, $z \in H$ if and only if $x, y \in H$. For S a polyhedral cone and $z \in \operatorname{int}(R_{x,y}) \subset S$, $z \in H$ if and only if $x, y \in H$.*

Proof. Write H as $\{w \in \mathbb{R}^m \,|\, \mathcal{L}(w) = b\}$ with $\mathcal{L} : \mathbb{R}^m \to \mathbb{R}$ a linear form and $b \in \mathbb{R}$. H supports S, so we lose no generality assuming $\mathcal{L}(S) \subset (-\infty, b]$.

For the polytope case, suppose $\mathcal{L}(x) = \mathcal{L}(y) = b$. For any $z \in (x, y)$, $z = \lambda x + (1 - \lambda)y$ with $0 < \lambda < 1$ and then

$$\mathcal{L}(z) = \lambda b + (1 - \lambda)b = b.$$

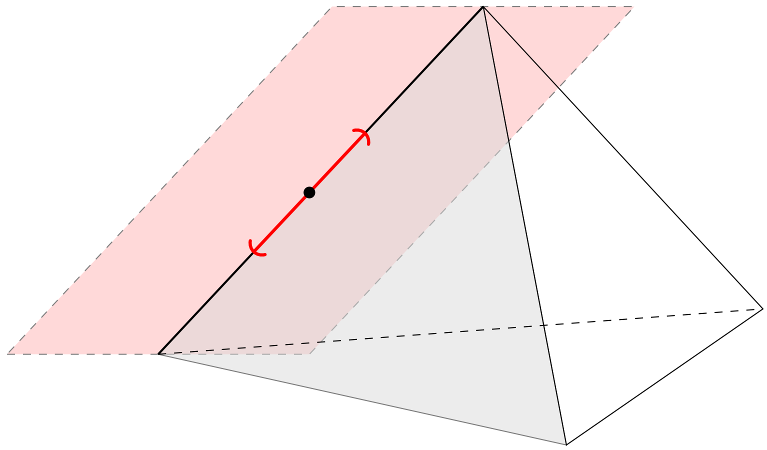


Figure 1.1: Example of Proposition 1.2.17.

For polyhedral cone case, $b = 0$ allows taking any linear combination of x and y.

Conversely, for polytopes, suppose $z = \lambda x + (1-\lambda)y$, $0 < \lambda < 1$, and $\mathcal{L}(z) = b$. Since

$$\mathcal{L}(z) = \lambda\mathcal{L}(x) + (1-\lambda)\mathcal{L}(y) = b,$$

either $b \in (\mathcal{L}(x), \mathcal{L}(y))$ or $\mathcal{L}(x) = \mathcal{L}(y) = b$, but $\mathcal{L}(x), \mathcal{L}(y) \leq b$ rules out the former. The polyhedral cone case follows similarly. $\square$

CHAPTER 2

The Metric Cone

In this chapter we give a brief introduction to the geometry of the metric cone, along with a short review of the important literature relevant to the work of this thesis. Towards the end of this section we give the construction of a new class of extreme graphical rays on the metric cone.

2.1 Definitions and Immediate Consequences

Definition 2.1.1. A *pseudometric* on X is a function $d : X \times X \to [0, \infty)$ such that for any $x, y, z \in X$,

1. $d(x, x) = 0$
2. $d(x, y) = d(y, x)$ (Symmetry)
3. $d(x, z) \leq d(x, y) + d(y, z)$ (Triangle Inequality)

We do not impose the positive-definiteness which characterizes metric. So pseudometrics allow for zero distances between distinct points. For $x, y, z \in X$, should $d(x, z) = d(x, y) + d(y, z)$, we call the triangle inequality *tight* (or *degenerate*).

We denote the set of pseudometrics on a set X by $\mathcal{M}(X)$. For the remainder of this chapter, we abuse terminology and abbreviate pseudometric to metric. Write $\mathbf{n} := \{1, \ldots, n\}$ and $\mathcal{M}_n := \mathcal{M}(\mathbf{n})$. For each pair of points (i, j) with $i < j$ we assign the distance $d_{ij} := d(i, j)$. We thus regard $\mathcal{M}_n$ as a subset of $\mathbb{R}^{\binom{n}{2}}$, identify metric d with vector $(d_{12}, d_{13}, \ldots, d_{1n}, d_{23}, \ldots, d_{n-1,n}) \in \mathbb{R}^{\binom{n}{2}}$, components listed in dictionary order.

Theorem 2.1.2 (Avis,1980)**.** *$\mathcal{M}_n$ forms an $\binom{n}{2}$-dimensional pointed polyhedral cone in $\mathbb{R}^{\binom{n}{2}}$ with facets given by*

$$d_{ij} + d_{jk} - d_{ik} = 0 \quad \textit{for} \quad 1 \leq i, j, k \leq n$$

Proof. First we prove $\mathcal{M}_n$ is a cone. We appeal to the half-space definition of a cone. It is clear that for each triple (i, j, k), $d_{ij} + d_{jk} - d_{ik} \leq 0$ defines a closed half-space in $\mathbb{R}^{\binom{n}{2}}$ whose boundary passes through the origin. Since $\mathcal{M}_n$ equals the intersection of finitely many such closed half-spaces, it forms a cone in $\mathbb{R}^{\binom{n}{2}}$. A small ball around point $(1, \ldots, 1) \in \mathbb{R}^{\binom{n}{2}}$ lies in the interior of the cone, making $\mathcal{M}_n$ $\binom{n}{2}$-dimensional. To see that equations

$$d_{ij} + d_{jk} - d_{ik} = 0 \quad \text{for} \quad 1 \leq i, j, k \leq n$$

give the facets of $\mathcal{M}_n$, consider a hyperplane H that intersects the boundary of $\mathcal{M}_n$, but not its interior. H must intersect the boundaries of some closed half-spaces given by triangle inequalities. Such an intersection will have maximum dimension when H intersects one single such closed half-space H_{ijk}, representing $d_{ij} + d_{jk} - d_{ik} = 0$.

To show that $H_{ijk} \cap \mathcal{M}_n$ forms a facet we must show that it has a non-empty interior with respect to the subspace topology. Equivalently, we must show that H_{ijk} contains a metric not in any other H_{lrs}. Take for d the metric with

$$d_{ij} = d_{jk} = 1/2, \quad d = 1 \quad \text{otherwise.}$$

This d lies in H_{ijk} but no other H_{lrs}. That makes $H_{ijk} \cap \mathcal{M}_n$ is a maximal face (or facet) of $\mathcal{M}_n$. □

Set $\Delta_{ijk}(d) := d_{ij} + d_{jk} - d_{ik}$. Say that metrics $d, \rho \in \mathcal{M}_n$ have the same *tight constraints* if $\Delta_{ijk}(\rho) = 0$ exactly when $\Delta_{ijk}(d) = 0$. Since facets correspond to tight constraints we have:

Corollary 2.1.3. *$d \in \mathcal{M}_n$ generates an extreme ray of $\mathcal{M}_n$ if and only if $\rho \in \mathcal{M}_n$ has the* *same tight constraints as d implies $\rho = \lambda \cdot d$.*

2.2 Extreme Rays

Henceforth, abusing terminology, we call a metric d an extreme ray of $\mathcal{M}_n$ if d generates one. The characterization of extreme rays of metric cones has received some attention. The most basic extreme rays arise from bi-partitions (or splits) of the set $\mathbf{n}$.

Definition 2.2.1. For disjoint, non-empty A and B with $A \cup B = \mathbf{n}$ (a split of $\mathbf{n}$), we call

$$\delta_{A,B}(x,y) := \begin{cases} 0 & \text{if } x,y \in A \text{ or } x,y \in B \\ 1 & \text{otherwise} \end{cases}$$

a *split* metric.

Proposition 2.2.2. *$\delta_{A,B}$ is an extreme ray of $\mathcal{M}_n$.*

Given a class of extreme metrics, we can ask what types of metrics live in their conical span. For the split extreme metrics, the conical span comprises precisely the so-called tree metrics, metrics characterized by the following four point condition:

$$d_T(i,j) + d_T(k,l) \leq \max\{d_T(i,k) + d_T(j,l), d_T(j,k) + d_T(i,l)\}.$$

Buneman [8] gives a complete decomposition of tree metrics into splits. The name arises because every non-negatively weighted tree T on n points defines a tree metric d_T.

Theorem 2.2.3 (Buneman,1970)**.** *Every metric d_T associated to a weighted tree T can be expressed in the form*

$$d_T = \sum_{\textit{splits}} \alpha_{A,B}\delta_{A,B}$$

where $\alpha_{A,B} = \frac{1}{2} \max_{\substack{i,j \in A \\ k,l \in B}} \{d_T(i,k) + d_T(j,l) - d_T(i,j) - d_T(k,l)\}$.

Bendelt and Dress [10] generalized Buneman's results to give a canonical decomposition for symmetric functions (and hence metrics) into splits and a split-prime component. They define the isolation indices $\alpha^d_{A,B}$ for a metric d and splits (A, B).

$$\alpha^d_{A,B} = \frac{1}{2} \min_{\substack{i,j \in A \\ k,l \in B}} \{\max \{d(i,k) + d(j,l), d(i,l) + d(j,k), d(i,j) + d(k,l)\} - d_{ij} - d_{kl}\}.$$

These indices were first derived for metrics on 4 points, and then generalized to give the result:

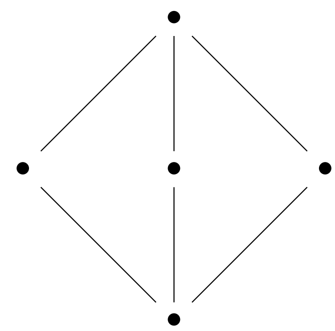
Theorem 2.2.4 (Bandelt-Dress, 1992). *Every symmetric function $d : X \times X \to \mathbb{R}$ on a finite set X can be expressed in the form*

$$d = d_0 + \sum_{d-splits} \alpha^d_{A,B} \delta_{A,B}$$

where d_0 is split-prime (d_0 has no d_0-splits). If d is a metric then so is d_0.

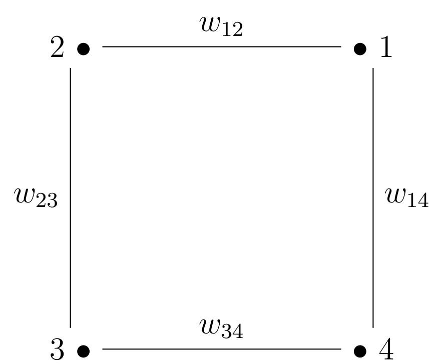
In the field of phylogeny, one associates dissimilarity coefficients to pairs of species via differences in their morphology. From these coefficients one can construct a metric and build a phylogenetic tree dervied from the split component of the metric in the decomposition above.

A *split-prime metric* is a metric which cannot be written in terms of splits; so outside their conical span. The first example of an extreme split-prime metric occurs in $\mathcal{M}_5$ represented as graphical metric by the graph $K_{2,3}$. We compute distances in a graph by taking the minimum length of a path between two vertices.



$K_{2,3}$

To prove the extremality of $K_{2,3}$, we first prove a lemma Avis [1] introduced. Suppose a positively weighted 4-cycle C_4



induces a metric, with distances w_{ij} for $1 \leq i < j \leq 4$, satisfying the tight constraints

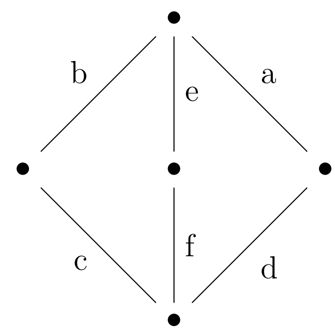
$$w_{12} + w_{23} = w_{13}$$
$$w_{14} + w_{34} = w_{13}$$
$$w_{12} + w_{14} = w_{24}$$
$$w_{23} + w_{34} = w_{24}.$$

Since solving this system by elimination yields $w_{12} = w_{34}$ and $w_{23} = w_{14}$, we have

Lemma 2.2.5 (Avis, 1980)**.** *Let (G, w) be a weighted undirected graph that contains a 4-cycle (i, j, k, l) having the same tight constraints as C_4, then $w_{ij} = w_{kl}$ and $w_{jk} = w_{il}$.*

Corollary 2.2.6. *$K_{2,3}$ is an extreme ray of $\mathcal{M}_5$*

Proof. Let d be a metric with the same tight constraints as $K_{2,3}$ then d can be represented by the weighted graph



By Lemma (2.2.5) $e = c$, $b = f$, $a = f$, and $e = d$. Similarly, $a = c$ and $d = b$. Hence, we obtain the 6-cycle of equalities $(a\, f\, b\, d\, e\, c)$, so that $d = a \cdot K_{2,3}$. □

Corollary 2.2.7. *The complete bipartite graph $K_{2,n-2}$ is an extreme ray of $\mathcal{M}_n$.*

$K_{2,3}$ is an example of a graphical extreme ray. Using the above lemma Avis [1] was able to show extremality for a large class of graphical metrics. First a definition.

Given a set of vertices V in a graph G, write $|V|$ for its cardinality and $||V||$ for the number of edges in the subgraph it induces.

Definition 2.2.8. A *dense m-partite graph G* is graph in which the vertex set can be partitioned into disjoint sets $V_1, \ldots, V_m$ with properties

1. $|V_m| \geq |V_{m-1}| \geq \cdots \geq |V_1| \geq 3$

2. $||V_i \cup V_j|| \geq |V_i| \cdot |V_j| - \max\{|V_i|, |V_j|\} + 2$ for $1 \leq i < j \leq m$

(Observe that $||V_i \cup V_j|| \leq |V_i| \cdot |V_j|$.)

Theorem 2.2.9 (Avis, 1980)**.** *If G is a dense m-partite graph of order n then d_G is an extreme ray of M_n*

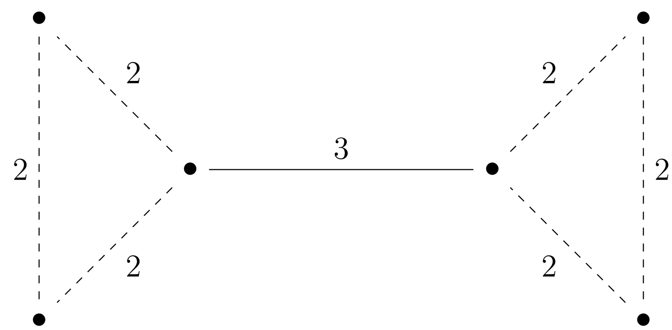
From this result Avis showed that almost all graphs on n points of "medium" density are extreme rays, and that extreme rays can have arbitrary local structure:

Theorem 2.2.10 (Avis, 1980)**.** *If d is a rational metric on m points, there exists an n and an extreme ray $\rho \in \mathcal{M}_n$ so that $\rho|_m = d$.*

2.3 Bowtie Metrics

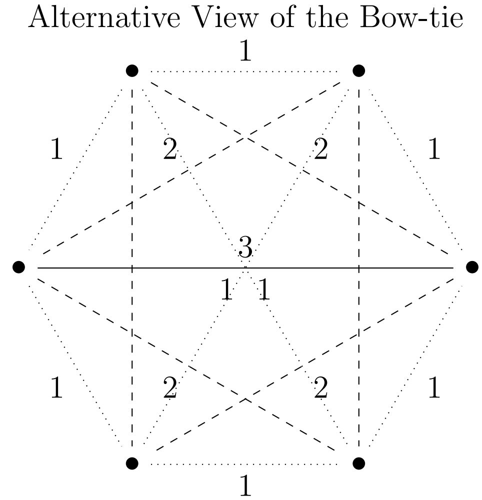
Inspired by Avis, we derive a new class of graphical metrics sitting outside his class of dense m-partite graphs. We start by considering what we call the *bow-tie* metric $\mathcal{B}$ on 6 points, pictured below; all unspecified distances equal 1.

The Bow-tie



We find the following alternative hexagonal view of the Bow-tie suggestive.

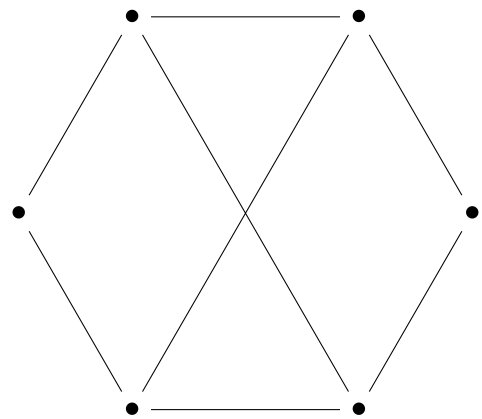
Alternative View of the Bow-tie



solid = 3, dashed = 2, dotted = 1

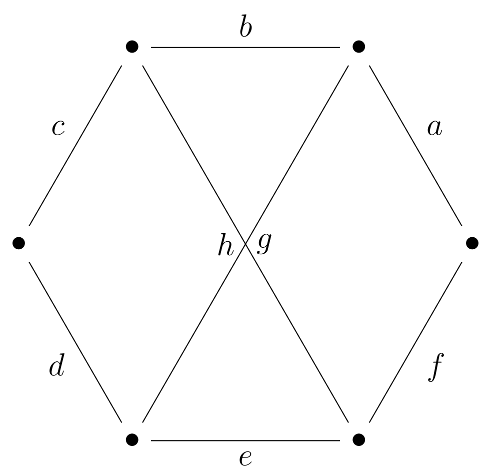
We can also represent the bow-tie metric graphically:

Graphical View of the Bow-tie



Proposition 2.3.1. *The bow-tie metric $\mathcal{B}$ is an extreme ray of $\mathcal{M}_6$.*

Proof. Consider a metric d having all of the same tight constraints as the bow-tie. We have labelled selective d distances as shown.

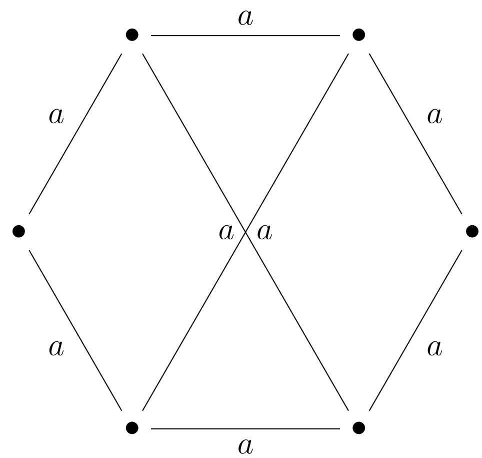


The 4-cycle with edge lengths $\{a, b, g, f\}$, say, has the same tight constraints as C_4, so Avis' lemma gives $b = f$ and $a = g$. Similarly considering all the visible 4-cycles gives, altogether,

$$b = f, \quad b = d, \quad c = e, \quad a = e$$

$$a = g, \quad c = h, \quad d = h, \quad f = g,$$

yielding the 8-cycle $(f\,b\,d\,h\,c\,e\,a\,g)$ of equalities. That means d coincides with the graphical metric



and thus $d = a \cdot \mathcal{B}$, making the bow-tie an extreme ray of $\mathcal{M}_6$. □

The symmetric structure of the bow-tie suggests a generalization to any number of points $n \geq 6$. First we need some definitions. Write Π_n for the *partitions* of $\mathbf{n}$.

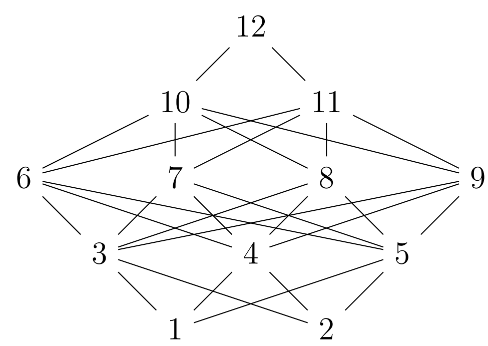
Definition 2.3.2. Call $\pi = \{P_1, \ldots, P_k\} \in \Pi_n$ a *cut partition* if $\max P_i \leq \min P_{i+1}$ for all $1 \leq i < k$.

$\pi = \{\{1,2\}, \{3,4,5\}, \{6\}\}$ is a cut partition of $\mathbf{6}$, for example. The cells of a partition enjoy an obvious linear order inherited from $\mathbf{n}$. The sequence of sizes of its cells determine the cut partition, allowing us to notate this π, say, as $2,3,1$.

Definition 2.3.3. Given a cut partition $\pi = \{P_1, \ldots, P_k\}$, let graph G_π have edges connecting every x in P_i to every x' in P_{i+1}, for i from 1 to $k-1$. Write $\mathcal{B}_\pi \in \mathcal{M}_n$ for the graphical metric associated to G_π.

Henceforth we add the size restrictions $|P_i| \geq 2$ for $2 \leq i \leq k-1$, $|P_1|, |P_k| \geq 1$

The original bow-tie metric $\mathcal{B}$ thus equals $\mathcal{B}_{1,2,2,2,1}$. For a more complex example consider the metric associated to graph $G_{2,3,4,2,1}$.



$G_{2,3,4,2,1}$

Lemma 2.3.4 (Kehoe, 2019)**.** *$\mathcal{B}_{l,r,s,w}$ constitutes an extreme ray of $\mathcal{M}_{l+r+s+w}$.*

Proof. Suppose metric d has the same tight constraints as $\mathcal{B}_{l,r,s,t}$. Regard d as the weighted graph $(G_{l,r,s,t}, w)$ for some non-negative weight function w on the edges of $G_{l,r,s,t}$.

$G_{l,r,s,t}$ generally contains many copies of the bowtie $G_{1,2,2,1}$ and any two edges that fall in a common bowtie have equal weights.

Two edges connecting vertices of P_2 and P_3 have either three or four endpoints between them. Choosing a vertex in P_1, a vertex in P_4, and only if necessary, another vertex from either P_2 or P_3 as needed, we get six vertices that induce a bowtie. All its edges must have equal weight, in particular, the two with which we started. So all edges connecting vertices of P_2 and P_3 have the equal weights.

But all edges connecting either P_1 and P_2, or P_3 and P_4, extend to vertex induced bowties that must indeed include edges connecting P_2 and P_3. So *all* edges have the same weight. w constant on all of G guarantees the extremality of $\mathcal{B}_{l,r,s,t}$. □

Theorem 2.3.5 (Kehoe, 2019)**.** *Let $n = n_1 + \cdots + n_k$ with $n_2, \ldots, n_{k-1} \geq 2$, and $n_1, n_k \geq 1$ and $k \geq 4$. Then $\mathcal{B}_{n_1,\ldots,n_k}$ constitutes an extreme ray of $\mathcal{M}_n$. In addition, $\mathcal{B}_{n_1,\ldots,n_k}$ exhibits as distances all integer values from 1 to $k-1$.*

Proof. Let $G_{n_1,\ldots,n_k}$ be the graph associated to the graphical metric $\mathcal{B}_{n_1,\ldots,n_k}$. Suppose that d is another metric which satisfies the same tight constraints as $\mathcal{B}_{n_1,\ldots,n_k}$. Then we can regard d as the weighted graph $(G_{n_1,\ldots,n_k}, w)$ for some nonnegative weight function on the edges of

$G_{n_1,\ldots,n_k}$. Consider the cut partition $\pi = \{P_1, \ldots, P_k\}$ associated to $\mathcal{B}_{n_1,\ldots,n_k}$ and define G_i as the restriction of the graph $G_{n_1,\ldots,n_k}$ to the set of vertices $P_i \cup P_{i+1} \cup P_{i+2} \cup P_{i+3}$ for $1 \leq i \leq k-3$. Then by construction G_i is isomorphic to the graph associated to the generalized bow-tie $\mathcal{B}_{n_i,n_{i+1},n_{i+2},n_{i+3}}$. By Lemma (2.3.4) $w|_{G_i}$ is constant. Since $G_i \cap G_{i+1} = P_{i+1} \cup P_{i+2} \cup P_{i+3}$ for $1 \leq i \leq k-3$, we have that w is in fact constant over all of G and hence $\mathcal{B}_{n_1,\ldots,n_k}$ is an extreme ray.

To show the second half of the theorem, consider any path from P_1 to P_k. Then every integer distance between 1 and $k-1$ will be achieved by $\mathcal{B}_{n_1,\ldots,n_k}$ along that path. □

Corollary 2.3.6 (Kehoe, 2019)**.** *For each $n \geq 5$ there are graphical extreme rays $d \in \mathcal{M}_n$ with rational distances $\left\{{}^1\!/_q, {}^2\!/_q, \ldots, 1\right\}$ for $2 \leq q \leq \lfloor \frac{n}{2} \rfloor$.*

Proof. For $n = 5$ just take the extreme ray $\frac{1}{2} \cdot K_{2,3}$. Now let $n \geq 6$ and consider first n even. Let $\pi = \{P_1, \ldots, P_k\}$ be the cut partition with $|P_1|, |P_k| = 1$ and $|P_i| = 2$ otherwise. Notice that $k = \lfloor \frac{n}{2} \rfloor + 1$. If we let $n_i = |P_i|$ then $\left(\frac{1}{k-1}\right) \cdot \mathcal{B}_{n_1,\ldots,n_k}$ is an extreme ray that has rational distances $\left\{{}^1\!/_q, {}^2\!/_q, \ldots, 1\right\}$ where $q = \lfloor \frac{n}{2} \rfloor$. We can obtain all other rational distances with denominator $3 \leq q < \lfloor \frac{n}{2} \rfloor$ by just considering the metrics $\mathcal{B}_{n_1,\ldots,n_{k-2},n_{k-1}+n_k}$, $\mathcal{B}_{n_1,\ldots,n_{k-3},n_{k-2}+n_{k-1}+n_k}, \ldots, \mathcal{B}_{n_1,n_2,n_3+\cdots+n_k}$. Now to obtain an extreme ray with denominator 2 on $n \geq 6$ points, we simply scale the bipartite graph $K_{2,n-2}$ by a half.

For odd n is odd, simply require instead that $|P_k| = 2$ in our initial choice of cut partition.

□

CHAPTER 3

The Metric Body

Whereas the set of pseudometrics form a convex cone, bounding metrics in the cone so no distance equals more than 1 yields a convex polytope. By considering this truncated cone, extreme rays give rise to extreme points, but many more arise. Focusing on this polytope, we can utilize linear programming to study both this new object and the cone, both theoretically and experimentally.

In this chapter we first start out by introducing the reader to the elements of bounded-by-1 pseudo metrics and there associated edge graphs. We give a complete characterization of extremality for an important class of extreme metrics, the so-called half-one metrics. We use this characterization to give a decomposition of the upper half of the polytope, and finally leave off with an interesting conjecture on the local geometry of extreme half-ones.

3.1 Definitions and Immediate Consequences

Let X be an arbitrary set.

Definition 3.1.1. A *bounded-by-1* pseudometric on X is a function $d : X \times X \to [0, 1]$ such that for any $x, y, z \in X$,

1. $d(x, x) = 0$
2. $d(x, y) = d(y, x)$ (Symmetric)
3. $d(x, z) \leq d(x, y) + d(y, z)$ (Triangle Inequality)

For a specific choice of $x, y \in X$ we'll call bounding constraint $d(x,y) \leq 1$ *tight* (or *unital*), in the case that equality holds.

We define the set $\bar{\mathcal{M}}(X)$ as the set of bounded-by-1 pseudometrics on X and write $\bar{\mathcal{M}}_n$ for $\bar{\mathcal{M}}(\mathbf{n})$. Analogous to the cone case, we have:

Proposition 3.1.2. *$\bar{\mathcal{M}}_n$ forms an $\binom{n}{2}$-dimensional polytope in $\mathbb{R}^{\binom{n}{2}}$ with facets given by*

$$d_{ij} + d_{jk} - d_{ik} = 0 \quad \textit{for} \quad 1 \leq i, j, k \leq n$$
$$d_{ij} = 1 \quad \textit{for} \quad 1 \leq i, j \leq n$$

Proof. We appeal to the half-space definition of a polytope. We can express $\bar{\mathcal{M}}_n$ as the intersection

$$\bar{\mathcal{M}}_n = \mathcal{M}_n \cap [0,1]^{\binom{n}{2}}.$$

We get an H-rep $\bar{\mathcal{M}}_n$ from the H-rep of $\mathcal{M}_n$ by adding some of the half-spaces connected to the obvious H-rep of the compact unit cube $[0,1]^{\binom{n}{2}}$. (The half-spaces bounded by hyperplanes passing through the origin add no new information.) This makes $\bar{\mathcal{M}}_n$ a polytope in $\mathbb{R}^{\binom{n}{2}}$. $\bar{\mathcal{M}}_n$ contains an open ball around the point $(\frac{1}{2}, \ldots, \frac{1}{2})$ in its interior, making $\bar{\mathcal{M}}_n$ $\binom{n}{2}$-dimensional.

As with the cone, equations

$$d_{ij} + d_{jk} - d_{ik} = 0 \quad \text{for} \quad 1 \leq i, j, k \leq n$$

produce facets of $\bar{\mathcal{M}}_n$. To see that $d_{ij} = 1$ for $1 \leq i, j \leq n$ form facets of $\bar{\mathcal{M}}_n$, we exhibit a bounded metric $d \in \mathcal{M}_n$ with exactly one unital bounding constraint and no degenerate triangle inequalities: make all distances between distinct points, save one, equal to $\frac{3}{4}$, and set the the remaining distance 1. □

Corollary 3.1.3. *Let $d \in \bar{\mathcal{M}}_n$. Then d is an extreme point of $\bar{\mathcal{M}}_n$ if and only if, for any*

$\rho \in \bar{\mathcal{M}}_n$ with the same tight constraints as d, we have that $\rho = d$. Thus an extreme metric is completely determined by its degenerate triangle inequalities and unital bounding constraints.

In all that follows, metric means a member of $d \in \bar{\mathcal{M}}(X)$, and we refer to $\bar{\mathcal{M}}_n$ as the *metric body* (for n points). We wish to understand, to the extent possible, the extreme points of $\bar{\mathcal{M}}(X)$. Firstly, as already noted, every extreme ray of $\mathcal{M}_n$ generates an extreme point of $\bar{\mathcal{M}}_n$.

Theorem 3.1.4. *An extreme ray $d \in \mathcal{M}_n$ gives rise to an extreme point $\left(\frac{1}{\max(d)}\right) \cdot d \in \bar{\mathcal{M}}_n$.*

Proof. Suppose d generates an extreme ray in $\mathcal{M}_n$. Then d has at least one non-zero distance, making $\frac{1}{\max(d)}$ well defined. Suppose d attains its maximum at d_{ij}. Defining $\tilde{d} = \left(\frac{1}{\max(d)}\right) \cdot d$, $\tilde{d} \in \bar{\mathcal{M}}_n$ and $\tilde{d}_{ij} = 1$. Now suppose $\rho \in \bar{\mathcal{M}}_n$ is another metric which satisfies the same tight constraints as $\tilde{d}$. Then since d is an extreme ray we must have that $\rho = \lambda \cdot d$ for some $\lambda > 0$, and hence $\rho_{ij} = \max(\rho)$. Since ρ must also share the same length 1 edges with $\tilde{d}$, we have that $\rho_{ij} = 1$. Thus $1 = \lambda \cdot \max(d)$, so that $\rho = \tilde{d}$. Hence, $\tilde{d}$ is an extreme point. □

As we will see shortly, extreme points of the body generally don't generate extreme ray of the cone. The basic idea is that if we start with an extreme bounded metric and drop the bounding constraints, perturbations may arise that distort the original length 1 edges independently. Moving from the body to the cone, the equality of length 1 edges becomes the equality of edges, so new constraints that cut down the dimension of the cone and alter the set of extreme rays.

Define the bounded generalized bow-tie metrics by

$$\overline{\mathcal{B}}_{n_1,\ldots,n_k} := \left(\frac{1}{\max(\mathcal{B}_{n_1,\ldots,n_k})}\right) \cdot \mathcal{B}_{n_1,\ldots,n_k}.$$

From our previous results we get

Corollary 3.1.5 (Kehoe, 2019)**.** *$\overline{\mathcal{B}}_{n_1,\ldots,n_k}$ is an extreme point of $\bar{\mathcal{M}}_{n_1+\cdots+n_k}$. In addition, $\overline{\mathcal{B}}_{n_1,\ldots,n_k}$ attains all rational values from $\left\{\frac{1}{k-1}, \frac{2}{k-1}, \ldots, \frac{k-2}{k-1}, 1\right\}$.*

Corollary 3.1.6 (Kehoe, 2019)**.** *For each $n \geq 5$ there are extreme metrics $d \in \bar{\mathcal{M}}_n$ with rational distances $\left\{ {}^{1}\!/_{q}, {}^{2}\!/_{q}, \ldots, 1 \right\}$ for $2 \leq q \leq \lfloor \frac{n}{2} \rfloor$.*

3.2 Separable Metric Spaces

On may consider metrics and bounded metrics on sets of any cardinality and the notions of extreme ray and extreme point carry over. In this section we apply our knowledge of finite metrics to exhibit some wild behavior in the general situation: modulo a countable number of points, any separable metric space with distances bounded by 1 occurs as an extreme metric.

Definition 3.2.1. We call a metric space (X, d) *separable* if X contains a countable dense subset.

Regardless of the cardinality of X, $\bar{\mathcal{M}}(X)$ constitutes a compact convex set. An *extreme metric* d (i.e. extreme point $d \in \bar{\mathcal{M}}(X)$) means d doesn't fall in the interior of any line segment contecting two points of $\bar{\mathcal{M}}(X)$. A *perturbation* of d means a symmetric function $\varepsilon : X \times X \to \mathbb{R}$ such that $(d - \varepsilon, d + \varepsilon) \subset \bar{\mathcal{M}}(X)$.

Given the tight constraint

$$d(x, y) + d(y, z) = d(x, z)$$

for $d \in \bar{\mathcal{M}}(X)$, Corollary (1.3.11) implies that

$$\varepsilon(x, y) + \varepsilon(y, z) = \varepsilon(x, z).$$

Thus if any one of the quantities $\varepsilon(x, y)$, $\varepsilon(y, z)$, or $\varepsilon(x, z)$ does not equal 0, then so does at least one other. In other words, perturbing one edge length in a tight triangle entails perturbing at least two lengths. In particular, perturbing $d(x, z)$, say, would force one also to perturb either $d(x, y)$ or $d(y, z)$ or both.

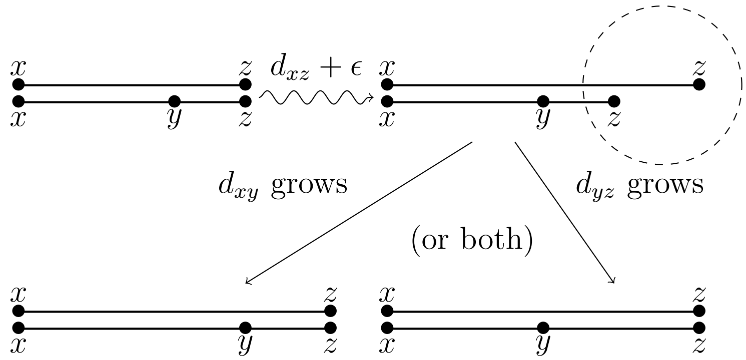


Figure 3.1: Effect of a Perturbation

Asserting that d admits only the zero perturbation characterizes d as extreme.

Definition 3.2.2. A metric space (Y, ρ) *extends* a metric space (X, d) if $X \subset Y$ and $\rho|_X = d$.

Theorem 3.2.3 (Kehoe, 2019)**.** *Every (X, d) separable metric space can extend to an extreme separable metric space $(\tilde{X}, \tilde{d})$ with $\tilde{X} \setminus X$ countable.*

The proof depends on two lemmas.

Lemma 3.2.4. *Given metric spaces (X, d), (Y, ρ) with $X \cap Y$ finite and non-empty, and with $d|_{X\cap Y} = \rho|_{X\cap Y}$, there exists a metric ω on $X \cup Y$ restricting to both d and ρ on X and Y respectively, and such that any distance $\omega(x, y)$ with $x \in X \setminus Y$ and $y \in Y \setminus X$ satisfies either $\omega(x, y) = 1$ or, for some $z \in X \cap Y$,*

$$\omega(x, y) = \omega(x, z) + \omega(z, y) = d(x, z) + \rho(z, y).$$

Proof. Defining $\omega|_X = d$ and $\omega|_Y = \rho$, leaves defining ω between points of X and Y not in the intersection. For $x \in X \setminus Y$ and $y \in Y \setminus X$ then define

$$\omega(x, y) := \min\left\{\min_{z\in X\cap Y} \{d(x, z) + \rho(z, y)\}, 1\right\}.$$

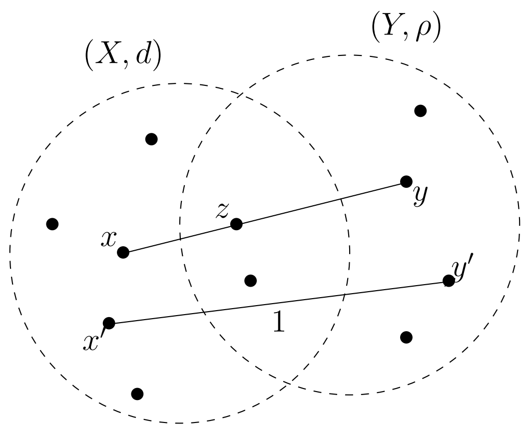


Figure 3.2: Definition of ω

Checking that ω defines a metric reduces to checking the triangle inequality for triples intersecting both X and Y. By symmetry it suffices to check the case where $x, z \in X$ and $y \in Y$. Suppose first that $z \in X \cap Y$ then we check the following three inequalities

1. $\omega(x,y) \leq \omega(x,z) + \omega(z,y)$

$$\begin{aligned}\omega(x,y) &\leq \min_{w \in X \cap Y} \{d(x,w) + \rho(w,y)\} \\ &\leq d(x,z) + \rho(z,y) \\ &= \omega(x,z) + \omega(z,y)\end{aligned}$$

2. $\omega(x,z) \leq \omega(x,y) + \omega(y,z)$

$$\begin{aligned}\omega(x,z) &\leq \min\left\{\min_{w \in X \cap Y} \{d(x,w) + d(w,z)\}, 1\right\} \\ &= \min\left\{\min_{w \in X \cap Y} \{d(x,w) + \rho(w,z)\}, 1\right\} \\ &\leq \min\left\{\min_{w \in X \cap Y} \{d(x,w) + \rho(w,y) + \rho(y,z)\}, 1\right\} \\ &\leq \min\left\{\min_{w \in X \cap Y} \{d(x,w) + \rho(w,y)\}, 1\right\} + \rho(y,z) \\ &= \omega(x,y) + \omega(y,z)\end{aligned}$$

3. $\omega(y,z) \leq \omega(x,y) + \omega(x,z)$.

   Follow the previous proof with the roles of x and y reversed.

We now consider $x, z \in X \setminus Y$ and $y \in Y \setminus X$, and check the three triangle inequalites

1. $\omega(x,y) \leq \omega(x,z) + \omega(z,y)$

$$\begin{aligned}\omega(x,y) &= \min\left\{\min_{w\in X\cap Y}\{d(x,w)+\rho(w,y)\},1\right\}\\&\leq \min\left\{\min_{w\in X\cap Y}\{d(z,w)+d(x,z)+\rho(w,y)\},1\right\}\\&\leq \omega(x,z)+\omega(z,y)\end{aligned}$$

2. $\omega(x,z) \leq \omega(x,y) + \omega(y,z)$

$$\begin{aligned}\omega(x,z) &\leq \min\left\{d(x,z)+2\cdot\min_{w\in X\cap Y}\{\rho(w,y)\},1\right\}\\&\leq \min\left\{\min_{w\in X\cap Y}\{d(x,w)+d(z,w)+\rho(w,y)+\rho(w,y)\},1\right\}\\&\leq \min\left\{\min_{w\in X\cap Y}\{d(x,w)+\rho(w,y)\},1\right\}\\&\quad+\min\left\{\min_{u\in X\cap Y}\{d(z,u)+\rho(u,y)\},1\right\}\\&= \omega(x,y)+\omega(y,z)\end{aligned}$$

3. $\omega(y,z) \leq \omega(x,y)+\omega(x,z)$. Follow the previous proof with the roles of x and y reversed.

Thus ω defines a metric on $X \cup Y$. By construction, for any $x \in X \setminus Y$ and $y \in Y \setminus X$ we have that $\omega(x,y)$ satisfies

$$\begin{aligned}\omega(x,y) &= \omega(x,z)+\omega(z,y)\\&= d(x,z)+\rho(z,y)\end{aligned}$$

for some $z \in X \cap Y$, or $\omega(x,y) = 1$. $\square$

Lemma 3.2.5. *Let (X, d) be a metric space with some distance $d(x_0, y_0)$ irrational. There exists a countable set $C = \{x_1, x_2, \ldots\}$ and a metric ω on $X \cup C$ such that the following properties are satisfied*

1. $\omega|_X = d$

2. $\omega(x_i, x_{i+1}) > \omega(x_{i+1}, x_{i+2})$ *for* $i \geq 0$

3. $\omega(x_i, x_j) = \sum_{k=i}^{j-1} \omega(x_k, x_{k+1}) \quad$ *for* $\quad i < j$

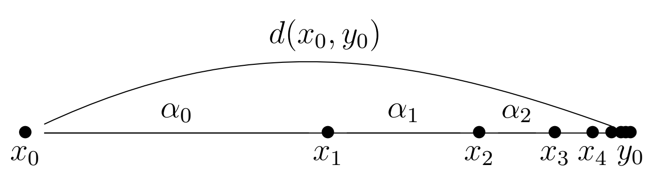
4. $\sum_{i=0}^{\infty} \omega(x_i, x_{i+1}) = \omega(x_0, y_0) = d(x_0, y_0)$

5. $\omega(x_j, y_0) = d(x_0, y_0) - \omega(x_0, x_j) = \sum_{i=j}^{\infty} \omega(x_i, x_{i+1})$

6. $\omega(x, x_i) = \min\{\omega(x, x_0) + \omega(x_0, x_i), 1\}$ *for* $x \in X \setminus \{y_0\}$

7. *If* $\omega \in (\omega_0, \omega_1)$ *then*
$$\sum_{i=0}^{\infty} \omega_0(x_i, x_{i+1}) = \omega_0(x_0, y_0)$$
and
$$\sum_{i=0}^{\infty} \omega_1(x_i, x_{i+1}) = \omega_1(x_0, y_0)$$
In other words $\sum_{i=0}^{\infty} \omega(x_i, x_{i+1}) = \omega(x_0, y_0)$ *is a (generalized) tight constraint.*

Proof. Use the decimal expansion of irrational $d(x_0, y_0)$ to write $d(x_0, y_0) = \sum_{i=0}^{\infty} \alpha_i$ with (α_i) a strictly decreasing non-negative sequence of rational numbers approaching 0 (roughly exponentially). View the set C, consisting of partial sums of the series together with 0 and $d(x_0, y_0)$, as points in a copy of the reals. Let C inherit a metric from the reals.



Identify $0, d(x_0, y_0) \in C$ with x_0 and y_0 respectively. Finally, as per the previous lemma, combine the two metrics, on X and on C to obtain a metric ω on $X \cup C$.

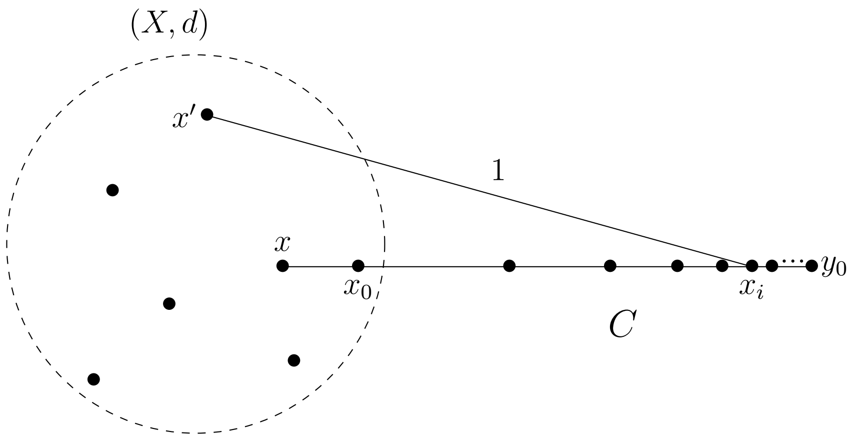


Figure 3.3: Definition of ω (note $y_0 \in X$)

ω defines a metric on $\tilde{X} := X \cup C$ satisfying properties 1 through 6 in the hypothesis of the lemma.

To prove condition 7 we will first define a sequence of linear functionals on the vector space $V_{\tilde{X}}$ of real-valued symmetric functions on $\tilde{X} \times \tilde{X}$. Let $\mathcal{L}_j : V_{\tilde{X}} \to \mathbb{R}$ be the linear functional defined as

$$\mathcal{L}_j(\eta) = \left(\sum_{i=0}^{j-1} \eta(x_i, x_{i+1}) \right) - \eta(x_0, x_j)$$

for $j \geq 1$. Now if $\eta \in \mathcal{M}(\tilde{X})$, by the triangle inequality we have that $\mathcal{L}_j(\eta) \geq 0$ for all j. Suppose now that $\omega_0, \omega_1 \in \mathcal{M}(\tilde{X})$ and $\omega \in (\omega_0, \omega_1)$. So $\omega = \lambda\omega_0 + (1 - \lambda)\omega_1$ for some $0 < \lambda < 1$. Since $\mathcal{L}_j(\omega) = 0$ for all j, by Proposition (1.3.11)

$$\mathcal{L}_j(\omega_0) = \mathcal{L}_j(\omega_1) = 0 \quad (*)$$

for all $j \geq 1$ and hence both $(\omega_0(x_0, x_j))$ and $(\omega_1(x_0, x_j))$ define increasing sequences of real numbers. Now, since

$$\omega(x_0, x_j) + \omega(x_j, y_0) = d(x_0, y_0) = \omega(x_0, y_0) \quad (**)$$

for any $j \geq 1$ we have that ω_0 and ω_1 must satisfy these same tight constraints. This yields us upper bounds

$$\omega_0(x_0, x_j) \leq \omega_0(x_0, y_0) \quad , \quad \omega_1(x_0, x_j) \leq \omega_1(x_0, y_0)$$

for all $j \geq 1$. Since bounded monotonic sequences converge we have that $\lim_{j\to\infty} \omega_0(x_0, x_j)$ and $\lim_{j\to\infty} \omega_1(x_0, x_j)$ exist. Using $(*)$ we obtain their values

$$\lim_{j\to\infty} \omega_0(x_0, x_j) = \sum_{i=0}^{\infty} \omega_0(x_i, x_{i+1}) \quad (***)$$

and

$$\lim_{j\to\infty} \omega_1(x_0, x_j) = \sum_{i=0}^{\infty} \omega_1(x_i, x_{i+1}) \quad (***).$$

By $(**)$ we also have the limits $\lim_{j\to\infty} \omega_0(x_j, y_0)$ and $\lim_{j\to\infty} \omega_1(x_j, y_0)$ exist. Now, since $\omega \in (\omega_0, \omega_1)$, in particular we have

$$\omega_0(x_j, y_0) < \omega(x_j, y_0) < \omega_1(x_j, y_0).$$

Thus $0 \leq \lim_{j\to\infty} \omega_0(x_j, y_0) \leq \lim_{j\to\infty} \omega(x_j, y_0) = 0$, so that

$$\lim_{j\to\infty} \omega_0(x_j, y_0) = 0.$$

We can compute the corresponding limit for ω_1 as

$$\begin{aligned}\lim_{j\to\infty} \omega_1(x_j, y_0) &= \lim_{j\to\infty} \frac{\omega(x_j, y_0) - \lambda\omega_0(x_j, y_0)}{1-\lambda} \\ &= 0.\end{aligned}$$

From this and $(**)$ we calculate the limits $\lim_{j\to\infty} \omega_0(x_0, x_j)$ and $\lim_{j\to\infty} \omega_1(x_0, x_j)$ as

$$\lim_{j\to\infty} \omega_0(x_0, x_j) = \omega_0(x_0, y_0)$$

and

$$\lim_{j\to\infty} \omega_1(x_0, x_j) = \omega_1(x_0, y_0)$$

Combining these equations with $(***)$ we obtain

$$\sum_{i=0}^{\infty} \omega_0(x_i, x_{i+1}) = \omega_0(x_0, y_0)$$

and

$$\sum_{i=0}^{\infty} \omega_1(x_i, x_{i+1}) = \omega_1(x_0, y_0)$$

□

We now prove our theorem.

Proof of Theorem (3.2.3). Since (X, d) is separable there exists a dense countable subset $A \subset X$. We will show that we can rigidify any distance in A by adding only at most countably many points to X. We will break the proof down as follows.

1. Take each irrational distance in A and divide it into countably many rational distances using Lemma (3.2.5).

2. For each rational distance in A or rational distance generated by an irrational distance of A, fit this rational distance into an extreme metric on a finite set. Use Lemma (3.2.4) to extend the metric d to include these extreme metrics.

3. After dividing irrational distances and rigidifying rational distances, we have extended the metric on X and only added countably many points. Call this new larger metric space $(\tilde{X}, \tilde{d})$ with corresponding dense set $\tilde{A}$.

4. By construction perturbing any non-unital distance in $\tilde{A}$ will have the effect of perturbing some non-unital rational distance in $\tilde{A}$. Every rational distance will be rigid and hence no distance in $\tilde{A}$ can move.

5. Since $\tilde{A}$ is dense in $\tilde{X}$ no distance in $\tilde{X}$ can move.

We first show that we can divide all the irrational distances into countably many rational distances. Define the set

$$Z = \{\{x, y\} \subset A \,|\, d(x, y) \text{ is irrational}\}$$

Since A is countable, Z is also countable. Hence we can order the pairs in Z as $\{\{x_1, y_1\}, \{x_2, y_2\}, \ldots\}$. By Lemma (3.2.5) there exists a countable set $C_1 = \{x_{1,1}, x_{1,2}, \ldots\}$ and metric ω_1 on $X \cup C_1$ extending d such that properties 1 through 6 in Lemma (3.2.5) are satisfied with $x_{1,0} := x_1$ and

$$\sum_{i=0}^{\infty} \omega_1(x_{1,i}, x_{1,i+1}) = d(x_1, y_1).$$

Using recursion, there exists a countable set $C_n = \{x_{n,1}, x_{n,2}, \ldots\}$ and a metric ω_n on $X_n := X \cup C_1 \cup \cdots \cup C_n$ extending ω_{n-1} satisfying the properties of Lemma (3.2.5) with $x_{n,0} := x_n$ and

$$\sum_{i=0}^{\infty} \omega_n(x_{1,i}, x_{1,i+1}) = d(x_n, y_n).$$

for $n \geq 1$. Now let

$$X_\infty = \bigcup_{n \geq 1} X_n$$

and define a metric ω on X_∞ by the property that $\omega|_{X_n} = \omega_n$. Note, since ω_{n+1} is an extension of ω_n for all $n \geq 1$ we have that ω is well defined. Similarly define $A_n := A \cup C_1 \cup \cdots \cup C_n$ and

$$A_\infty = \bigcup_{n \geq 1} A_n.$$

We claim that if we perturb any non-unital edge length in A_n, this must have the effect of perturbing a non-unital rational edge length in A_n or a non-unital edge length in A . We proceed by induction on n, taking as a basis for induction A_1. Let $\omega_1(x, y) < 1$ be any non-unital edge length in A_1 and suppose we perturb this edge length. We then have three cases to consider

1. $x \in A \setminus \{y_1\}$ and $y \in C_1$.

   In this case $y = x_{1,i}$ for some $i \geq 1$ so that

   $$\omega_1(x, y) = d(x, x_1) + \omega_1(x_1, x_{1,i})$$

   since this constraint is tight, perturbing $\omega_1(x, y)$ will have the effect of either perturbing $d(x, x_1)$, a non-unital edge length in A, or perturbing $\omega_1(x_1, x_{1,i})$, a non-unital rational length in A_1.

2. $x = y_1$ and $y \in C_1$

   In this case $y = x_{1,i}$ for some $i \geq 1$ so that

   $$\omega_1(x, y) = d(x_1, y_1) - \omega_1(x_1, x_{1,i}).$$

   Since this constraint is tight, perturbing $\omega_1(x, y)$ will have the effect of either perturbing $d(x_1, y_1)$, a non-unital edge length in A, or perturbing $\omega_1(x_1, x_{1,i})$, a non-unital rational length in A_1.

3. $x, y \in C_1$

   In this case $\omega_1(x, y)$ already constitutes a rational edge length in A_1.

This completes the basis for induction. Now assume the induction hypothesis holds for all $k < n$. Let $\omega_n(x, y)$ be any non-unital edge length in A_n. Without loss of generality we can

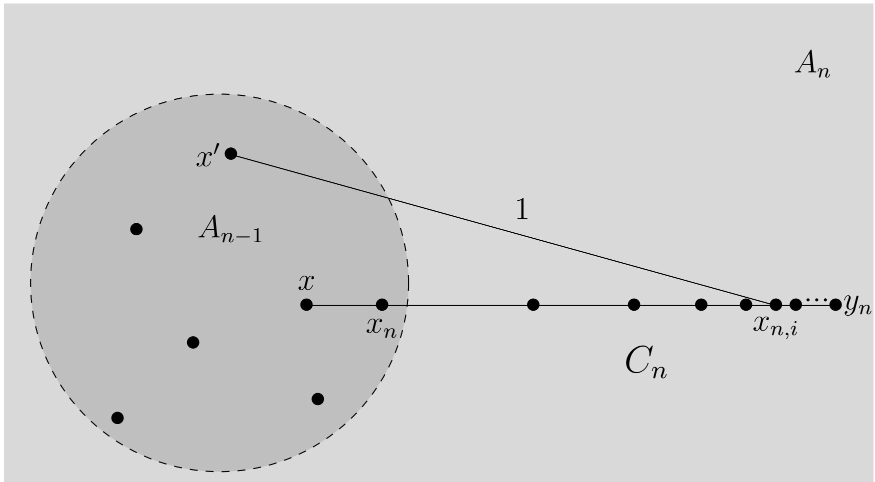


Figure 3.4

assume that x and y are not both in A_{n-1}. We have the following similar cases

1. $x \in A_{n-1} \setminus \{y_n\}$ and $y \in C_n$.

   In this case $y = x_{n,i}$ for some $i \geq 1$ so that

$$\omega_n(x, y) = \omega_{n-1}(x, x_n) + \omega_n(x_n, x_{n,i})$$

   since this constraint is tight, perturbing $\omega_n(x, y)$ will have the effect of either perturbing $\omega_{n-1}(x, x_1)$, or perturbing $\omega_n(x_n, x_{n,i})$, a non-unital rational length in A_n. By the induction hypothesis, perturbing $\omega_{n-1}(x, x_1)$ will have the effect of either perturbing a non-unital rational edge length in $A_{n-1} \subset A_n$ or perturbing a non-unital edge length in A.

2. $x = y_n$ and $y \in C_n$.

   Similar to case 1.

3. $x, y \in C_n$

In this case $\omega_n(x, y)$ already constitutes a non-unital rational edge length in A_n.

This completes the induction.

Perturbing any non-unital edge length in A_∞ thus entails perturbing a non-unital rational edge length in A_∞. Indeed, any irrational edge length $d(x_i, y_i)$ in A must satisfy

$$d(x_i, y_i) = \sum_{j=0}^{\infty} \omega_i(x_{i,j}, x_{i,j+1})$$

which Lemma (3.2.5) makes a tight constraint. Hence, perturbing $d(x_i, y_i)$ entails perturbing some non-unital rational edge length $\omega_i(x_{i,k}, x_{i,k+1})$ in $A_i \subset A_\infty$ with $k \geq 0$.

We now show that we can rigidify every positive, non-unital rational distance in A_∞. First define the set

$$W = \{\{x, y\} \subset A_\infty \,|\, d(x, y) \in \mathbb{Q} \cap (0, 1)\}$$

Since A_∞ is countable, W is also countable. Hence we can order pairs in W as $\{\{x_1, y_1\}, \{x_2, y_2\}, \ldots\}$. First consider the rational distance $\omega(x_1, y_1)$ in A_∞. By Corollary (3.1.6) there exists an $n_1 > 0$ and an extreme metric $\rho_1 \in \mathcal{M}_{n_1}$ such that $\omega(x_1, y_1)$ is a value of ρ_1. Let (i_0, j_0) be the pair of points in $\mathbf{n_1}$ where $\rho_1(i_0, j_0) = \omega(x_1, y_1)$. By taking the disjoint union of X_∞ and $\mathbf{n_1}$, and then identifying the points x_1 and y_1 with i_0 and j_0 respectively we obtain a new set $\tilde{X}_1$ where we view X_∞ and $\mathbf{n_1}$ as subsets. By Lemma (3.2.4) there exists a metric $\tilde{d}_1$ on $\tilde{X}_1$ that restricts to both ω and ρ_1 on X_∞ and $\mathbf{n_1}$ respectively, and such that any distance $\tilde{d}_1(x, y)$ with $x \in X_\infty \setminus \mathbf{n_1}$ and $y \in \mathbf{n_1} \setminus X_\infty$ satisfies

$$\begin{aligned} \tilde{d}_1(x, y) &= \tilde{d}_1(x, z) + \tilde{d}_1(z, y) \\ &= \omega(x, z) + \rho_1(z, y) \end{aligned}$$

for some $z \in \{x_1, y_1\}$, or $\tilde{d}_1(x, y) = 1$. Since $\tilde{d}_1(z, y)$ lies in an extreme metric, we cannot perturb it. Now define the subset $\tilde{A}_1 := A_\infty \cup \mathbf{n_1}$. Notice that perturbing non-unital $\tilde{d}_1(x, y)$

in $\tilde{A}_1$ with $x \in A_\infty \setminus \mathbf{n}$ and $y \in \mathbf{n} \setminus A_\infty$ entails perturbing a non-unital edge length in A_∞, and therefore entails perturbing a non-unital rational edge length in A_∞. By successively rigidifying the positive non-unital rational distances in A_∞ and extending the metric using Lemma (3.2.4) we obtain a sequence of metric spaces $\left((\tilde{d}_1, \tilde{X}_1), (\tilde{d}_2, \tilde{X}_2), \dots, (\tilde{d}_n, \tilde{X}_n), \dots\right)$ such that $(\tilde{X}_{i+1}, \tilde{d}_{i+1})$ is an extension of $(\tilde{X}_i, \tilde{d}_i)$ for all $i \geq 1$. Let

$$\tilde{X} = \bigcup_{n \geq 1} \tilde{X}_n$$

and define the metric $\tilde{d}$ on $\tilde{X}$ by the property $\tilde{d}|_{\tilde{X}_n} = \tilde{d}_n$. Define the subset of $\tilde{X}$

$$\tilde{A} = A \cup \left(\tilde{X} \setminus X\right).$$

By an induction similar to the above, it is simple to show that perturbing any non-unital edge length in $\tilde{A}$ will have the effect of perturbing a non-unital rational edge length in A_∞. But every non-unital rational edge length in A_∞ lies in an extreme metric, and hence cannot be perturbed. Thus, no edge length in $\tilde{A}$ can be perturbed and we have $\tilde{d}|_{\tilde{A}}$ extreme.

Finally, we show that $(\tilde{X}, \tilde{d})$ is an extreme separable extension of (X, d). Since $(\tilde{X}, \tilde{d})$ was built from (X, d) using successive extensions it follows that $(\tilde{X}, \tilde{d})$ is an extension of (X, d). The set $\tilde{A}$ will be a dense subset of $\tilde{X}$. Indeed, if $x \in \tilde{X}$ then either $x \in X$ or $x \in \tilde{X} \setminus X$. In the latter case, $x \in \tilde{A}$. In the former case, any $\tilde{d}$-neighborhood of x will undoubtedly contain a d-neighborhood of x. Since A is dense in X, this means any $\tilde{d}$-neighborhood of x must contain points from $A \subset \tilde{A}$. Hence, $\tilde{A}$ is dense in $\tilde{X}$. Since $\tilde{d}|_{\tilde{A}}$ is extreme and $\tilde{d}$ is completely determined by its values on a dense set, it follows that $\tilde{d}$ is also extreme. This proves the theorem. □

3.3 Extreme Metrics

As noted, bi-partitions give extreme points in the metric body. Actually any partition of the underlying set X corresponds to an extreme point in the body. Call metric $d \in \bar{\mathcal{M}}(X)$ *discrete* if it takes only the values zero and one.

Associate a partition of X with any (pseudo)metric on X via the equivalence relation $\sim_d$:

$$x \sim_d y \iff d(x, y) = 0 \text{ for } x, y \in X$$

With discrete metrics, we lose no information when we pass to its partition, so henceforth we identify the set of discrete metrics on X with the set $\Pi(X)$ of partitions on X.

Proposition 3.3.1. *Partitions give extreme points:* $\Pi(X) \subset ex(\bar{\mathcal{M}}(X))$.

Proof. Let $d \in \Pi(X)$ and assume (for contradiction) that there exists a non-constant line segment $d_t \in \bar{\mathcal{M}}(X)$, $t \in [0, 1]$, with $d = d_{t_*}$ for some $t_* \in (0, 1)$.

For $x, y \in X$, suppose that $d(x, y) = 0$. Then

$$d_{t_*}(x, y) = t_* d_1(x, y) + (1 - t_*) d_0(x, y) = 0.$$

But $t_*, 1 - t_* > 0$ and $d_1(x, y), d_0(x, y) \geq 0$. Thus $d_1(x, y), d_0(x, y) = 0$ and $d_t(x, y) = 0$ for all t.

Now suppose instead that $d(x, y) = 1$. Then

$$d_{t_*}(x, y) = t_* d_1(x, y) + (1 - t_*) d_0(x, y) = 1.$$

We can't have $d_1(x,y) < 1$ lest

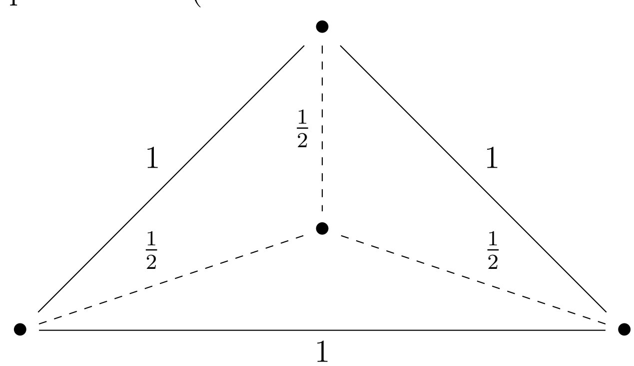
$$\begin{aligned} d_{t_*}(x,y) &< t_* + (1-t_*)d_0(x,y) \\ &\leq t_* + 1 - t_* \\ &= 1 \end{aligned}$$

So $d_1(x,y) = 1$, and similarly $d_0(x,y) = 1$, and so $d_t(x,y) = 1$ for all t.

Having d_t constant contradicts the assumption. □

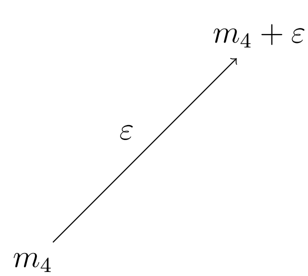
We first encounter a non-partition extreme metric on 4 points. This metric has a single point as the common "mid-point" of every edge in a triangle.

Midpoint Metric (Extreme Metric 1 from Table 4.1.1)



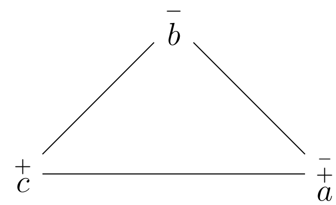
Proposition 3.3.2. *The mid-point metric m_4 is extreme*

Proof. Perturbation cannot affect length one edges when we consider metrics with distances bounded by 1. Label the half-length distances a, b, and c. Let $\varepsilon \in \mathbb{R}^{\binom{4}{2}}$ be a perturbation of m_4.



By the tightness of $a + b = 1$, perturbing distance a positively forces perturbing distance b negatively. But perturbing b negatively then forces perturbing c positively, which finally forces perturbing a negatively, contradiction. $\square$

We obtained an alternating sequence of positive/negative pertubations on the half-length edges of m_4. We can represent this as a cycle of signed nodes



Extremality rested on the impossibility of perturbing a (or any other length 1/2 edge here) both positively and negatively. This simple idea will generalize to a technique for proving extremality for large classes of metrics.

The next proposition allows restricting attention to positive-definite (i.e. true) metrics. Given a set X and a metric $d \in \bar{\mathcal{M}}(X)$, set $\overline{X}^d := X/_{\sim_d}$, the quotient by the equivalence relation $\sim_d$. Given two equivalence relations $\sim_1$ and $\sim_2$, say $\sim_1$ refines $\sim_2$ if $x \sim_1 y$ implies $x \sim_2 y$.

Proposition 3.3.3. *Fix a metric $d \in \bar{\mathcal{M}}(X)$ and equivalence relation $\sim$ on X with quotient $\overline{X}$. If $\sim$ refines $\sim_d$, then d descends to a metric $\bar{d}$ on $\overline{X}$. Moreover, the quotient by $\sim_d$ itself carries a true metric.*

Proof. $\bar{d}([x],[y]) = d(x,y)$ gives a well-defined metric. Indeed if $x \sim z$ and $y \sim w$, then $x \sim_d z$ and $y \sim_d w$, so $d(x,z) = 0$ and $d(y,w) = 0$. Then

$$d(x,y) \leq d(x,z) + d(z,y) = d(z,y)$$
$$d(z,y) \leq d(z,x) + d(x,y) = d(x,y)$$

and $d(x,y) = d(z,y)$, and similarly $d(x,y) = d(z,w)$. Certainly $\bar{d}$ inherits symmetry and takes its values in $[0,1]$ and if $\bar{d}$ violated the triangle inequality, so would d.

For $\sim_d$, $\bar{d}([x],[y]) = 0$ implies $d(x,y) = 0$ implies $x \sim_d y$ implies $[x] = [y]$ and positive-definiteness follows.

$\square$

Metrics also lift. Given $\rho \in \bar{\mathcal{M}}(\overline{X})$, define the *covering metric* on X by $\hat{\rho}(x,y) := \rho([x],[y])$. We need only check the triangle inequality.

$$\begin{aligned}\hat{\rho}(x,z) &= \rho([x],[z]) \\ &\leq \rho([x],[y]) + \rho([y],[z]) \\ &= \hat{\rho}(x,y) + \hat{\rho}(y,z)\end{aligned}$$

Note: $\bar{\hat{\rho}} = \rho$ for any ρ, $\hat{\bar{d}} = d$ for any refinement of $\sim_d$. So equal covers or (valid) quotients imply equal metrics.

An equivalence relation $\sim$ determines a set of metrics,

$$\bar{\mathcal{M}}^{\sim}(X) := \left\{d \in \bar{\mathcal{M}}(X) \middle| \sim \text{refines} \sim_d\right\}$$

$\bar{\mathcal{M}}^{\sim}(X)$ constitutes a convex subspace of $\bar{\mathcal{M}}(X)$. Indeed:

Proposition 3.3.4. *The map $\bar{\mathcal{M}}^{\sim}(X) \to \bar{\mathcal{M}}(\overline{X})$ given by $d \mapsto \bar{d}$ is a convex isomorphism. i.e a bijection that preserves convex combinations.*

Proof. Covering provides the inverse. $\square$

Theorem 3.3.5. *If $\sim$ refines $\sim_d$ and $d \in ex(\bar{\mathcal{M}}(X))$ then $\bar{d} \in ex(\bar{\mathcal{M}}(\overline{X}))$*

Proof. Suppose $d \in \mathrm{ex}(\bar{\mathcal{M}}(X))$ and suppose $\bar{d} \in [\rho_1, \rho_2]$ then $d = \hat{\bar{d}} \in [\hat{\rho}_1, \hat{\rho}_2]$. Thus $d = \hat{\rho}_1$

or $d = \hat{\rho}_2$. Without loss of generality assume $d = \hat{\rho}_1$, then we have,

$$\begin{aligned}\bar{d}([x],[y]) &= d(x,y)\\ &= \hat{\rho}_1(x,y)\\ &= \rho_1([x],[y])\end{aligned}$$

Thus $\bar{d} = \rho_1$ so that $\bar{d} \in \mathrm{ex}(\bar{\mathcal{M}}(\overline{X}))$ $\square$

Since extreme metrics descend to extreme metrics, any extreme metric d descends to an extreme true metric on a $\overline{X}^d$. Moreover, we can recover any metric as the cover of the unique true metric to which it descends.

Theorem 3.3.6. *Suppose $\rho \in ex(\bar{\mathcal{M}}(\overline{X}))$ then $\hat{\rho} \in ex(\bar{\mathcal{M}}(X))$*

Proof. Let$\rho \in \mathrm{ex}(\bar{\mathcal{M}}(\overline{X}))$. Suppose $\hat{\rho} \in [d_1, d_2]$ and that $\hat{\rho} = \lambda_1 d_1 + \lambda_2 d_2$ with $\lambda_1 + \lambda_2 = 1$ and $\lambda_1, \lambda_2 \geq 0$. Then d_1 and d_2 vanishes wherever $\hat{\rho}$ vanishes. That makes both $\bar{d}_1$ and $\bar{d}_2$ well-defined on $\overline{X}$. We then have $\rho = \lambda_1 \bar{d}_1 + \lambda_2 \bar{d}_2$, so that $\rho \in [\bar{d}_1, \bar{d}_2]$. ρ extreme implies $\rho = \bar{d}_1$ or $\rho = \bar{d}_2$. Say $\rho = \bar{d}_1$; then $\hat{\rho} = \hat{\bar{d}}_1 = d_1$. So $\hat{\rho} \in \mathrm{ex}(\bar{\mathcal{M}}(X))$. $\square$

Theorem 3.3.7. *The restricted map $\bar{\mathcal{M}}^{\sim}(X) \cap ex(\bar{\mathcal{M}}(X)) \to ex(\bar{\mathcal{M}}(\overline{X}))$ given by $d \mapsto \bar{d}$ is a bijection.*

Proof. Extreme metrics descend to extreme metrics and extreme metrics have extreme covers. Then we restrict the convex isomorphism given in Theorem 3.3.4 to get a bijection between $\bar{\mathcal{M}}^{\sim}(X) \cap \mathrm{ex}(\bar{\mathcal{M}}(X))$ and $\mathrm{ex}(\bar{\mathcal{M}}(\overline{X}))$. $\square$

All this makes classifying positive-definite extreme metrics tantamount to classifying all extreme metrics.

We now consider a simple type of co-product on metrics. The reader should understand the terminology as merely borrowed from category, seeing as we neither develop nor apply that perspective.

Definition 3.3.8. Given sets X and Y, and metrics $d \in \bar{\mathcal{M}}(X)$ and $\rho \in \bar{\mathcal{M}}(Y)$, define metric $d \sqcup \rho$ on $X \coprod Y$ by

$$d \sqcup \rho|_{X \times X} = d;$$

$$d \sqcup \rho|_{Y \times Y} = \rho;$$

$$d \sqcup \rho(x, y) = 1;$$

for $x \in X$ and $y \in Y$.

Proposition 3.3.9. *$d \sqcup \rho \in \bar{\mathcal{M}}(X \coprod Y)$. If $d \in ex(\bar{\mathcal{M}}(X))$ and $\rho \in ex(\bar{\mathcal{M}}(Y))$, then $d \sqcup \rho \in ex(\bar{\mathcal{M}}(X \coprod Y))$.*

Proof. A triangle not in just X or Y has two sides of length 1 and a third no longer, so the triangle inequality holds in general.

Given $d \in \mathrm{ex}(\bar{\mathcal{M}}(X))$ and $\rho \in \mathrm{ex}(\bar{\mathcal{M}}(Y))$, suppose $d \sqcup \rho = t\alpha + (1-t)\beta$ with $\alpha, \beta \in \bar{\mathcal{M}}(X \coprod Y)$ and $t \in (0, 1)$. Then α and β take the value 1 wherever $d \sqcup \rho$ does (see Theorem 3.3.1), so

$$\alpha = \alpha|_X \sqcup \alpha|_Y,$$

$$\beta = \beta|_X \sqcup \beta|_Y.$$

Then $d = t\alpha|_X + (1-t)\beta|_X$ and $\rho = t\alpha|_Y + (1-t)\beta|_Y$. Extremality of d and ρ then forces $\alpha|_X = \beta|_X$ and $\alpha|_Y = \beta|_Y$. So $\alpha = \beta$. ☐

Corollary 3.3.10 (Kehoe, 2019). *$\coprod_{i=1}^{n} \overline{\mathcal{B}}_{n_{i,1},\ldots,n_{i,k_i}}$ is an extreme point of $\bar{\mathcal{M}}_N$ where $N = \sum_{i,j} n_{i,j}$.*

Proposition 3.3.11. *Suppose d and ρ lie in the convex hull of $\{d_i\}$ and $\{\rho_j\}$ respectively. Then $d \sqcup \rho$ lies in the convex hull of $\{d_i \sqcup \rho_j\}$.*

Proof. First, for each j, express $d \sqcup \rho_j$ as a convex combination of metrics $d_i \sqcup \rho_j$. Then express $d \sqcup \rho$ as a convex combination of the $d \sqcup \rho_j$. ☐

Call a metric *irreducible* if it doesn't arise as a non-trivial co-product.

3.4 Geometric Structures

Given a set X, write $\mathcal{E}(X)$, the *edges* of X, for the set of unordered pairs $\{x, y\}$ and $\mathcal{T}(X)$, the *triangles* of X, for the set of unordered triples $\{x, y, z\}$. Symmetry allows considering a metric d as a function

$$d : \mathcal{E}(X) \to [0, 1].$$

Write $[x, y, z]$ for the triangle inequality

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

So we don't distinguish $[z, y, x]$ from $[x, y, z]$. Let $T(X)$ denote the set of such triples. When the context is clear, we'll omit the dependence on X for all geometric constructions on X. To maintain a consistency of notation with edges and triples, we'll often use the bracket notation $[x, y]$ to denote the edge $\{x, y\}$. Given two edges $[x, y]$ and $[y, z]$ sharing a common point we form the triple,

$$[x, y] \vee [y, z] := [x, y, z].$$

We consider that $[x, y]$ contained in $[x, y, z]$ and write

$$[x, y] \subset [x, y, z]$$

if $\{x, y\} \subset \{x, y, z\}$. We call an ordered triangle $[x, y, z]$ *degenerate*, with *long side* $[x, z]$, and *short sides* $[x, y]$ and $[y, z]$ if,

$$d(x, z) = d(x, y) + d(y, z).$$

Proposition 3.4.1. *A true metric $d \in \bar{\mathcal{M}}(X)$ has at most one degenerate triple with points* *in any given triangle* $\{x, y, z\} \in \mathcal{T}(X)$*.*

Given any distance function d and any degenerate ordered triangle, the other two triangle inequalities on the same points hold automatically.

Proof. Suppose, say, both $[x, y, z]$ and $[y, x, z]$, so $d(x, z) = d(x, y) + d(y, z)$ and also $d(y, z) = d(x, z) + d(x, y)$. Then $d(x, y) = d(x, z) - d(y, z) = d(y, z) - d(x, z) = 0$, and we don't have a true metric.

For any non-negative function d, $d(x, y) + d(y, z) = d(x, z)$ implies

$$d(y, z) + d(x, z) = 2d(y, z) + d(x, y) \geq d(x, y).$$

□

3.5 Metrics on a Finite Set and Linear Programming

In this section we introduce linear programming tools which enable us to classify a large elementary class of extreme metrics. Proposition (3.1.2) allows us to embed $\bar{\mathcal{M}}_n$ into $\mathbb{R}^m$ with $m = \binom{n}{2}$ using $\mathcal{E}(X)$ as a natural basis; we label coordinates of $\bar{\mathcal{M}}_n$ in dictionary order of the indices when convenient.

Once embedded, finding extreme metrics in $\bar{\mathcal{M}}_n$ translates to a linear programming problem. Specifically, $\bar{\mathcal{M}}_n$ produces a polytope in Euclidean space defined by the inequalities,

$$\begin{aligned}
d_{ik} - d_{ij} - d_{jk} &\leq 0, && 1 \leq i < j < k \leq n; && (3.5.1)\\
d_{kj} - d_{ki} - d_{ij} &\leq 0, && 1 \leq i < j < k \leq n; &&\\
d_{ji} - d_{jk} - d_{ki} &\leq 0, && 1 \leq i < j < k \leq n; &&\\
d_{ij} &\leq 1, && 1 \leq i < j \leq n; &&\\
d_{ij} &\geq 0, && 1 \leq i < j \leq n. && (3.5.2)
\end{aligned}$$

Thus $\bar{\mathcal{M}}_n$ equals the intersection of $3\binom{n}{3} + 2\binom{n}{2}$ half-spaces, written succinctly as $\bar{\mathcal{M}}_n = \{d \in \mathbb{R}^m \mid \quad Hd \leq b\}$.

We now develop some language to discuss our problem from the context of linear programming.

Definition 3.5.3. Call a vector $d \in \mathbb{R}^m$ *feasible* if it satisfies $Hd \leq b$.

In the context of metric spaces, feasibility means a vector that defines a metric.

Definition 3.5.4. Given a vector (not necessarily a metric) $d \in \mathbb{R}^m$, let A_d and a_d denote the maximum set of rows of H and b respectively so that $A_d \cdot d = a_d$. We call (A_d, a_d) the *active constraints*, A_d the *active matrix* and a_d the *active vector*. Call a vector $d \in \mathbb{R}^m$ a *basic solution* if A_d has a maximum rank, so (A_d, a_d) defines it.

The active constraints determine which supporting hyperplanes of $\bar{\mathcal{M}}_n$ in the H-rep contain d.

The theory of linear programming (LP) tells us that extreme points of a polytope coincide with basic feasible solutions of the associated LP-problem. Thus an extreme metric d will arise as the unique solution to an equation $Ad = a$ where A and a specify m equations refining the fundamental inequalities given (3.5.1), (3.5.2) above.

Every extreme metric d satisfies a defining system of linear *integer* equations, so $d \in \mathbb{Q}^m$. Thus we need only understand what makes rational metrics extreme. For a rational metric d let $\mathrm{den}(d)$ denote the minimal denominator (non-zero) such that $\mathrm{den}(d) \cdot d \in \mathbb{Z}^m$. If $\mathrm{den}(d) = r$ we'll call d an r-den metric. We'll write $\bar{\mathcal{M}}_n^r$ for all rational metrics with minimum denominator r and $\mathrm{ex}(\bar{\mathcal{M}}_n^r)$ for just the extreme ones.

Definition 3.5.5. Let $[i, j, k] \in T$ then we define $e_{[i,j,k]}$ to be the m-dimensional row vector such that,

$$
\begin{aligned}
e_{[i,j,k]}([i,k]) &= 1; \\
e_{[i,j,k]}([i,j]) &= -1; \\
e_{[i,j,k]}([j,k]) &= -1; \\
e_{[i,j,k]}([q,l]) &= 0 \ , \ \ [q,l] \ \nsubseteq [i,j,k].
\end{aligned}
$$

The row vector $e_{[i,j,k]}$ encodes the data from the $[i,j,k]$-triangle-inequality.

Definition 3.5.6. For $d \in \bar{\mathcal{M}}_n$,

$$\mathcal{U}_d = \{E \in \mathcal{E} | d_E = 1\},$$

comprises the *unital edges* associated to d, the edges on which d takes the value 1. $\mathcal{N}_d = \mathcal{E} \backslash \mathcal{U}_d$ comprises the *non-unital* edges.

Given a true metric $d \in \mathrm{ex}(\bar{\mathcal{M}}_n)$ with $r = |\mathcal{U}_d|$, we wish to write a minimal active matrix A which defines the extreme point d as a basic feasible solution to $Hd \leq b$ and retains all active unital constraints.

We begin with the tedious but necessary task of ordering. Take the edges $\mathcal{E}$ in dictionary order and identify $\mathcal{U}_d$ as a subset $\{i_q\}_{i=1}^{r}$ of $\{1, \ldots, m\}$ of cardinality r indexed so $i_q < i_{q+1}$. Similarly, we identify $\mathcal{N}_d$ as a subset $\{l_q\}_{i=1}^{m-r}$ of $\{1, \ldots, m\}$ of cardinality $m - r$ indexed so $l_q < l_{q+1}$. Write e_j for the the row vector in $\mathbb{R}^m$ with a 1 in the jth position, and 0 elsewhere.

Theorem 3.5.7 (Kehoe, 2017)**.** *d uniquely solves $Ad = a$ where (A_j denoting the jth row of A):*

$$
\begin{aligned}
A_{i_q} &= e_{i_q}, & 1 \leq q \leq r; \\
A_{l_q} &= e_{T_q}, & 1 \leq q \leq m - r;
\end{aligned}
$$

and

$$a_j = \begin{cases} 1 & j \in \{i_q\} \\ 0 & j \in \{l_q\}, \end{cases}$$

and the T_q represent distinct elements of $\mathcal{T}$ (not just of T), and each $E \in \mathcal{N}_d$ belongs to some T_q.

Proof. From the discussion earlier we know that an extreme true metric d will be the unique solution to an equation $Ad = a$ where A and a arise from m equations strengthening the inequalities (3.5.1), (3.5.2). Now let S denote the $r \times m$ matrix with rows taken from the set

$$\{e_{i_q} \mid 1 \leq q \leq r\}$$

Append the matrix S to the matrix A, use S to zero out any entries in columns corresponding to unital edges, and permute columns to bring non-unital edges to the left and unital edges to the right. Obtain thereby a block diagonal matrix

$$X := \begin{pmatrix} A' & 0 \\ 0 & S \end{pmatrix}.$$

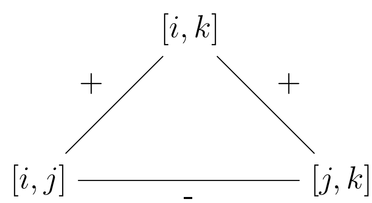
Since $\text{rank}(A') + \text{rank}(S) = m$ and $\text{rank}(S) = r$, we have that $\text{rank}(A') = m - r$. Thus we can modify A, substituting in rows of S while maintaining full rank and using row vectors corresponding to unital edges of d. Of course we must adjust the entries of a accordingly. Now for any $E \in \mathcal{N}_d$ there must be a degenerate triangle T containing E represented by a row in A, or else a column of A would equal 0. As we assume we have a true metric d, Theorem (3.4.1) says we have, represented in A, degenerate triangles distinct in $\mathcal{T}$. Finally obtain the desired A and a by now permuting rows so that they occur in an appropriate order, as per the requirements of the theorem. □

3.6 The Edge Graph

Start with Γ, the undirected graph whose nodes stand for the edges in $\mathcal{E}$, two nodes connected in Γ if the corresponding edges belong to a triangle. One calls Γ the *line graph* $L(K_n)$ of the complete graph K_n.

Given a metric $d \in \bar{\mathcal{M}}_n$ build a signed subgraph, Γ_d, of Γ, with nodes representing the non-unital edges in $\mathcal{N}_d$ and edges associated to degenerate triangles as follows.

Given a degenerate triangle of non-unital edges $[i, j, k]$ for d, (so long side $[i, k]$), connect the nodes representing $[i, j]$, $[j, k]$, and $[i, k]$ with positive and negative edges as depicted below.



Given a degenerate triangle $[i, j, k]$ for d with long side $[i, k]$ unital, simply connect $[i, j]$ and $[j, k]$ with a negative edge.

Γ_d encodes graphically the H-rep of d, allowing us to explore the linear dependence of degenerate triangles. As a signed graph, Γ_d has positive and negative induced subgraphs that we write as G^+ and G^-.

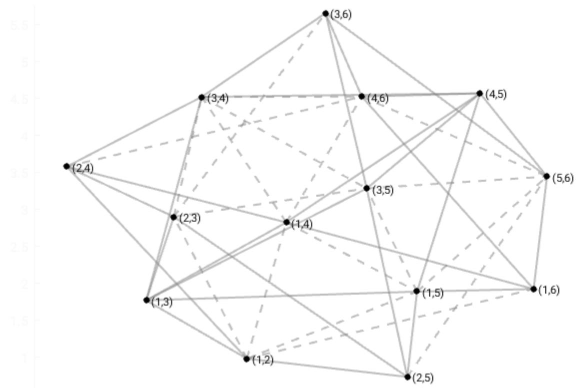


Figure 3.5: Edge Graph of the bow-tie metric $\overline{\mathcal{B}}_{1,2,2,1}$
(dashed=negative, solid=positive)

Definition 3.6.1. Call a signed subgraph κ of Γ short-sided if $\kappa = \kappa^-$.

The edges in short-sided graphs correspond to triangles in X, allowing the study of syzygies of degenerate triangles in terms of graph notions.

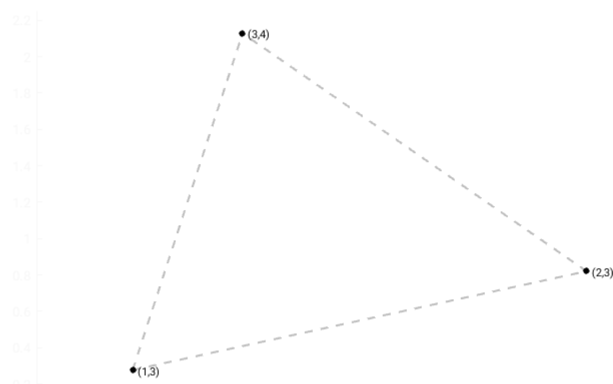


Figure 3.6: Edge Graph of the Midpoint Metric

Definition 3.6.2. By a *path* γ we shall mean a subgraph of Γ whose distinct nodes $\{[i_1, j_1], \ldots, [i_k, j_k]\}$ $(k \geq 2)$ occur connected each to the next and each consecutive pair $[i_l, j_l]$, $[i_{l+1}, j_{l+1}]$ share a common index. The edges of γ correspond to triangles $T_l = [i_l, j_l] \vee [i_{l+1}, j_{l+1}]$.

By the *edge space* of any subgraph $\kappa \subset \Gamma$ we mean the linear subspace $V_\kappa \subset \mathbb{R}^m$ (of dimension δ_κ) with vectors supported on the coordinates corresponding to vertices in κ. Let $\pi_\kappa : \mathbb{R}^m \to V_\kappa$ denote orthogonal projection.

Proposition 3.6.3. *Consider a short-sided path γ with vertices $\{[i_1, j_1], \dots, [i_k, j_k]\}$ and also the associated finite sequence $\{e_{T_l}\}_{l=1}^{k-1}$ of triangle inequality encoding vectors. Write γ_0 for the subpath $\{[i_2, j_2], \dots, [i_k, j_k]\}$ and V_{γ_0} for the vector space with basis the nodes of γ_0. Projecting the $\{e_{T_l}\}_{l=1}^{k-1}$ onto V_{γ_0} produces a linearly independent set, namely $\{\pi_{V_{\gamma_0}}(e_{T_l})\}_{l=1}^{k-1}$.*

Proof. Perform row reduction on the matrix having rows $\{\pi_{V_\gamma}(e_{T_l})\}_{l=1}^{k-1}$.

First replace e_{T_2} with, $R_2 = e_{T_1} - e_{T_2}$; so $R_2([i_1, j_1]) = -1, R_2([i_3, j_3]) = 1$ and zeros.

Next replace e_{T_3} with $R_3 = R_2 + e_{T_3}$, so $R_3([i_1, j_1]) = -1, R_3([i_4, j_4]) = -1$ and zeros.

In general $R_1 = e_{T_1}$ and e_{T_l} gets replaced with

$$R_l = R_{l-1} + (-1)^{l+1} e_{T_l} = \sum_{q=1}^{l} (-1)^{q+1} e_{T_q}.$$

Then $R_l([i_1, j_1]) = -1$ and $R_l([i_{l+1}, j_{l+1}]) = (-1)^l$ and zeros beside. Projection onto V_{γ_0} produces a diagonal matrix with non-zero diagonal entries, so linearly independent rows. Row operations don't affect linear independence, so the same holds for $\{\pi_{V_{\gamma_0}}(e_{T_l})\}_{l=1}^{k-1}$. □

For later use, we give the elementary matrix and inverse associated to the row operations in the proof:

$$E = \begin{pmatrix} 1 & & & \\ 1 & -1 & & \\ \vdots & \vdots & \ddots & \\ 1 & -1 & \dots & \pm 1 \end{pmatrix} \quad E^{-1} = \begin{pmatrix} 1 & & & & \\ 1 & -1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \mp 1 & \pm 1 \end{pmatrix}.$$

Corollary 3.6.4. *Let τ be a short-sided tree with root τ_*. If $\tau_0 = \tau \setminus \tau_*$ then $\{e_T\}_{T \in E(\tau)}$ is a linearly independent set over V_{τ_0}.*

Proof. Write τ as a union of short-sided paths $\tau = \gamma^1 \cup \dots \cup \gamma^p$ so that

1. γ^q has vertices $\left\{[i_1^q, j_1^q], \dots, [i_{k_q}^q, j_{k_q}^q]\right\}$;

2. $\tau_* = [i_1^1, j_1^1] \in \gamma^1$; and

3. $\gamma^q \cap (\gamma^1 \cup \cdots \cup \gamma^{q-1}) = [i_1^q, j_1^q]$ for all $1 < q \leq p$.

Induct on the number of paths in the union.

For the base case, apply (3.6.3) to $\tau = \gamma^1$.

For $\tau' = \tau \cup \gamma^s$, $\tau = \gamma^1 \cup \cdots \cup \gamma^{s-1}$, suppose, by induction, we already have linear independence for $\{e_T\}_{T \in E(\tau)}$ over V_{τ_0}.

Write $\tau'_0 = \tau' \setminus \tau_*$. Since $\gamma^s \cap \tau = [i_1^s, j_1^s]$, $V_{\gamma_0^s} \cap V_{\tau_0} = 0$. Then the linear independence of $\{e_T\}_{T \in E(\tau)}$ over V_{τ_0} and $\{e_T\}_{T \in E(\gamma^s)}$ over $V_{\gamma_0^s}$ together give linear independence of $\{e_T\}_{T \in E(\tau')}$ over $V_{\tau'_0} = V_{\gamma_0^s} \bigoplus V_{\tau_0}$, completing the induction. □

Definition 3.6.5. A *cycle* means a subgraph of Γ whose vertices we can list distinctly as $\{[i_1, j_1], \ldots, [i_k, j_k]\}$, $k \geq 3$, so that $[i_l, j_l]$ and $[i_{l+1}, j_{l+1}]$ share a common index, and also $[i_k, j_k]$ and $[i_1, j_1]$. The edges of the cycle connect $[i_l, j_l]$ to $[i_{l+1}, j_{l+1}]$ for $1 \leq j \leq k-1$, and $[i_k, j_k]$ to $[i_1, j_1]$. The cycle has associated triangles, first T_l for $1 \leq l \leq k-1$ same as the path from $[i_1, j_1]$ to $[i_k, j_k]$, and also $T_k := [i_k, j_k] \vee [i_1, j_1]$. Call a cycle *odd* or *even* according to the parity of k.

Proposition 3.6.6. *Fix a short-sided cycle θ with vertices $\{[i_1, j_1], \ldots, [i_k, j_k]\}$. $\{e_{T_l}\}_{l=1}^k$ forms a linearly independent set over V_θ if and only if θ has odd parity.*

Proof. Just modify the proof above for paths. Replace rows e_{T_l} for $1 \leq l \leq k-1$ with $R_l = \sum_{q=1}^{l} (-1)^{q+1} e_{T_q}$ getting values, $R_l([i_1, j_1]) = -1$ and $R_l([i_{l+1}, j_{l+1}]) = (-1)^l$ and zero otherwise. Then replace e_{T_k} with $R_k = \sum_{q=1}^{k} (-1)^{q+1} e_{T_q}$ getting values

$$R_k([i_1, j_1]) = \begin{cases} 0 & \text{k is even} \\ -2 & \text{k is odd} \end{cases}$$

and 0 otherwise.

k even makes $R_k = 0$ makes $\{R_1, \dots, R_{k-1}, R_k\}$ linearly dependent over V_θ, so likewise $\{e_{T_l}\}_{l=1}^k$. For k odd, replace R_j, $j < k$ with $R'_j = R_j - \frac{1}{2}R_k$. Then $\{R_k, R'_1, \dots, R'_{k-1}\}$ forms a diagonal matrix, hence the linearly independence of $\{e_{T_l}\}_{l=1}^k$ over V_θ. □

Again we can calculate the elementary matrix representing the final step in the row reduction above for the odd cycle,

$$C_k = \begin{pmatrix} 1 & \dots & \dots & -\frac{1}{2} & -\frac{1}{2} \\ & \ddots & & \vdots & \vdots \\ & & 1 & -\frac{1}{2} & -\frac{1}{2} \\ & & & \frac{1}{2} & -\frac{1}{2} \\ & & & 1 & 1 \end{pmatrix}, C_k{}^{-1} = \begin{pmatrix} 1 & \dots & \dots & \dots & \frac{1}{2} \\ & \ddots & & & \vdots \\ & & 1 & & \frac{1}{2} \\ & & & 1 & \frac{1}{2} \\ & & & -1 & \frac{1}{2} \end{pmatrix}$$

Theorem 3.6.7 (Kehoe, 2017)**.** *For $\kappa \subset \Gamma$ any connected short-sided subgraph with one odd cycle, we will have $\{e_T\}_{T \in E(\kappa)}$ is a linearly independent set over V_κ.*

Proof. Write $\kappa = \theta \cup \tau^1 \cup \dots \cup \tau^p$ with θ an odd cycle, the τ^q disjoint trees with each root $\tau^q_* \in \theta$. By construction θ and τ^q shares only the vertex τ^q_*. Proposition (3.6.6) gives $\{e_T\}_{T \in E(\theta)}$ linearly independent over V_θ, and Corollary (3.6.4) gives $\{e_T\}_{T \in E(\tau^q)}$ linearly independent set over $V_{\tau^q_0}$. Then from

$$\begin{aligned} V_\theta \cap V_{\tau^q_0} &= 0 & \qquad 1 \le q \le p \\ V_{\tau^q} \cap V_{\tau^s} &= 0 & \qquad q \ne s \end{aligned}$$

we have $\{e_T\}_{T \in E(\kappa)}$ linearly independent set over $V_\kappa = V_\theta \bigoplus V_{\tau^1_0} \bigoplus \dots \bigoplus V_{\tau^p_0}$.

□

We shall call a connected graph containing exactly one odd cycle a *germ* (think about what they look like), and a disjoint union of germs a *colony*. We now have the immediate corollary.

Corollary 3.6.8. *Given a short-sided colony $\kappa \subset \Gamma$, $\{e_T\}_{T \in E(\kappa)}$ forms a linearly independent set over V_κ.*

Theorem 3.6.9 (Kehoe, 2017)**.** *If Γ_d, for a true metric d, contains a spanning short-sided colony, then $d \in ex(\bar{\mathcal{M}}_n)$.*

Proof. Suppose Γ_d contains a spanning short-sided colony κ. Then $V_\kappa = \mathbb{R}^{m-r} \subset \mathbb{R}^m$ where $r = |\mathcal{U}_d|$, and $e(\kappa) = m - r$. Theorem (3.6.8) makes $\{e_T\}_{T \in E(\kappa)}$ linearly independent over $\mathbb{R}^{m-r}$. Label the $m - r$ triangles as T_q for $1 \leq q \leq m - r$ and define the matrix A and the vector a as,

$$A_{i_q} = e_{i_q} \qquad\qquad 1 \leq q \leq r$$

$$A_{l_q} = e_{T_q} \qquad\qquad 1 \leq q \leq m - r$$

with A_j denoting the jth row of A and,

$$a_j = \begin{cases} 1 & j \in \{i_q\} \\ 0 & j \in \{l_q\} \end{cases}$$

Then d uniquely solves $Ad = a$, so constitutes a basic feasible solution to the associated LP problem defined at the beginning of Section 3.5. Basic feasible solutions coincide with extreme points, so $d \in \mathrm{ex}(\bar{\mathcal{M}}_n)$. □

We now give a partial converse to Theorem (3.6.9).

Theorem 3.6.10 (Kehoe, 2017)**.** *Consider a true metric $d \in ex(\bar{\mathcal{M}}_n)$ with Γ_d short-sided. Then each component of Γ_d contains an odd cycle.*

Proof. Theorem (3.5.7) says d uniquely solves a type of matrix equation $Ad = a$. Writing A_j (resp. a_j) for the jth row of A (resp. a), A has the form

$$A_{i_q} = e_{i_q} \qquad 1 \le q \le r$$

$$A_{l_q} = e_{T_q} \qquad 1 \le q \le m - r,$$

and a the form

$$a_j = \begin{cases} 1 & j \in \{i_q\} \\ 0 & j \in \{l_q\} \end{cases}$$

Moreover we will have the T_q distinct in $\mathcal{T}$, and for each $E \in \mathcal{N}_d$ there will exist a q such that $E \subset T_q$. This last property guarantees that Γ_d spans.

Γ_d short-sided and spanning makes every long side of a triangle $T \in E(\Gamma_d)$ unital. Thus the triangles T_q associated to the vectors $A_{l_q} = e_{T_q}$ define a spanning subgraph κ of Γ_d with the property that $e(\kappa) = v(\kappa)$, as we must have the same number of equations and variables. Indeed, we argue that $e(\kappa^s) = v(\kappa^s)$ for each germ in the colony. Having $e(\kappa^s) > v(\kappa^s)$ would make $\{e_T\}_{T \in \kappa_s}$ linearly dependent over V_{κ_s}. By hypothesis, our degenerate triangles have their long edges unital, so this linear dependence would contradict A's having full rank.

Having κ^s connected with $e(\kappa^s) = v(\kappa^s)$ means κ^s contains only one cycle θ^s. If θ^s even, we get $\{e_T\}_{T \in E(\theta^s)}$ linearly dependent over V_{θ^s} by Theorem (3.6.6), contradicting full rank for A again. So we have θ^s odd. □

The question now arises, do extreme true metrics with short-sided spanning graphs Γ_d exist?

Theorem 3.6.11 (Kehoe, 2017)**.** *If Γ_d for a true metric d contains a spanning short-sided colony κ, d constitutes an extreme 2-den metric.*

Proof. Theorem (3.6.8) makes d extreme. Decompose $\kappa = \kappa^1 \sqcup \cdots \sqcup \kappa^p$ into germs. Theorem (3.5.7) then makes d the unique solution of matrix equation $Ad = a$ with the matrix A

associated to κ (including rows associated to unital edges). Permuting rows and columns (to group edges and triangle inequalities according to the germ that contains them) and row reduction techniques as in Theorems (3.6.6) and (3.6.3), we can write,

$$A = E(I_r \bigoplus N_1 \bigoplus \cdots \bigoplus N_p)$$

where $E \in M_m(\mathbb{Z})$, $|\det(E)| = 1$, I_r stands for the $r \times r$ identity matrix (corresponding to unital edges), and the N_q for $v(\kappa^q) \times v(\kappa^q)$ diagonal matrices having $|(N_q)_{11}| = 2$ and $|(N_q)_{ii}| = 1, i > 1$.

Then we have that $E^{-1} \in M_m(\mathbb{Z})$ and $N_q^{-1} = \frac{1}{2}M$ for $M \in M_{v(\kappa^q)}(\mathbb{Z})$ so that,

$$\begin{aligned} 2d &= 2A^{-1}a \\ &= 2(I_r \bigoplus N_1^{-1} \bigoplus \cdots \bigoplus N_p^{-1})E^{-1}a \\ &= (2I_r \bigoplus M_1 \bigoplus \cdots \bigoplus M_p)E^{-1}a \in \mathbb{Z}^m \end{aligned}$$

Using the hypotheses that $0 < d \leq 1$ we get that $d_I = \frac{1}{2}$ for any $I \in \mathcal{N}_d$. □

We end the chapter with two results characterizing extremality for 2-den true metrics or *half-one metrics*.

Corollary 3.6.12 (Kehoe, 2019)**.** *For a half-one metric d, extremality implies that each component of* Γ_d *contains an odd cycle and conversely. For* $d \in \mathcal{M}_n$ *an extreme ray half-one metric,* Γ_d *contains an odd cycle.*

Proof. For half-one metrics d, $\Gamma_d = \Gamma_d^-$. Then extreme rays in $\mathcal{M}_n$ give extreme points of $\bar{\mathcal{M}}_n$. □

Now one naturally asks what addition conditions on a half-one metric d, guarantee extremality for d *as a ray*? The simple nature of body extremality for half-one metrics raises hopes for a simple characterization of cone extremality. Getting such a large class of extreme rays might lead to a canonical decomposition of the split-prime component given in the decomposition by Bendelt and Dress (see Chapter 1).

Considering non-short-sided graphs might lead to insight into metrics with higher denominators. One might start with graphs containing only a single node incident to positive edges.

3.7 Perturbations

A non-extreme metric d must lie between two other metrics, neither equal to d itself. Say $d \in [d_\alpha, d_\beta]$; then both d_α and d_β share d's unital edges and degenerate triangles. Hence any perturbation $\varepsilon \in \mathbb{R}_m$ of d must satisfy,

1. $\varepsilon_{ij} + \varepsilon_{jk} = \varepsilon_{ik}$ whenever $d_{ij} + d_{jk} = d_{ik}$

2. $\varepsilon_{ij} = 0$ whenever $d_{ij} \in \{0, 1\}$

This identifies perturbations of d with the row space of A_d, P_d. Call a perturbation ε *feasible* when $d + \varepsilon$ constitutes a metric. Trivially

Proposition 3.7.1. *If $d_{ij} + d_{jk} = d_{ik} = 1$, then $\varepsilon_{ij} = -\varepsilon_{jk}$ for any perturbation ε.* □

Proposition 3.7.2. *Suppose metric d has connected short-sided edge graph Γ_d. Fix (τ, E_0), a rooted spanning tree for Γ_d. Denote tree distance from E_0 to E given by $l(E)$. Then any perturbation ε of d equals a constant multiple of,*

$$\eta(E) = \begin{cases} (-1)^{l(E)} & , E \in \mathcal{N}_d \\ 0 & , E \in \mathcal{U}_d \end{cases} \tag{3.7.3}$$

Proof. Γ_d short-sided makes every long side of a degenerate triangle unital. Thus we may identify degenerate triangles of d with the edges of Γ_d. Let ε be any perturbation of d and suppose E adjacent to E'. Proposition (3.7.1) says that $\varepsilon(E) = -\varepsilon(E')$. Then Γ_d connected makes ε determined by its value at the root of the tree E_0. Thus

$$\varepsilon(E) = \begin{cases} (-1)^{l(E)}\varepsilon(E_0) & , E \in \mathcal{N}_d \\ 0 & , E \in \mathcal{U}_d \ . \end{cases}$$

Hence $\varepsilon = \varepsilon(E_0)\eta$ as claimed. □

Corollary 3.7.4. *If metric d has short-sided edge graph Γ_d decomposed into components as $\Gamma_d = \kappa^1 \sqcup \cdots \sqcup \kappa^p$ and η_i defined, component by component, as in* (3.7.3)*, then any perturbation ε of d has the form*

$$\varepsilon(E) = \begin{cases} c_i\eta_i(E) & , E \in \kappa^i \\ 0 & , E \in \mathcal{U}_d \end{cases}$$

with the $c_i \in \mathbb{R}$. If κ_i contains an odd cycle, $c_i = 0$.

Definition 3.7.5. We call a component κ of Γ_d *rigid* if $\varepsilon|_\kappa = 0$ for every perturbation ε.

If d has a short-sided component κ, Proposition (3.6.6) makes κ rigid if and only if κ contains an odd cycle.

Corollary 3.7.6. *Let d be a metric with short-sided edge graph Γ_d and let D denote the number of non-rigid components of Γ_d. Then $\dim(P_d) = D$. For a basis for P_d, take $\{\eta_1, \ldots, \eta_D\}$, the η_i as defined in* (3.7.3)*.*

CHAPTER 4

Half-one Metrics

Let $\mathcal{H}_n$ and $\mathrm{ex}(\mathcal{H}_n)$ denote the set of half-one metrics and extreme half-one metrics on n points respectively. Any vector $d \in \mathbb{R}^m$ such that $d_i \in \left\{\frac{1}{2}, 1\right\}$ automatically gives a half-one metric on n points. When extreme r-den metrics compete for abundance, this automatic feasibility gives the half-one metrics and advantage. Half-ones also tend to have more (linear) dependencies among their degenerate triangles.

This chapter gives a lower bound for the cardinality of $\mathcal{H}_n$ and a new decomposition result for a significant part of the metric body.

4.1 A lower bound for $|\mathcal{H}_n|$

Given $d \in \mathcal{H}_n$, we build a graph G_d with vertices $\{1, \ldots, n\}$ and edges $\mathcal{N}_d$. So $d(x, y) = 1/2$ in d entails adjacency in G_d and conversely. G_d disconnected will say exactly that d equals the co-product of two half-one metrics on smaller sets. Every graph G has an associated line graph $L(G)$, vertices of $L(G)$ matching edges of G and connected when the matching edges share a vertex. A basic graph theory says passing from G to $L(G)$ preserves connectivity. Certainly, $\Gamma_d \subset L(G_d)$ but when does $L(G_d) = \Gamma_d$? The equality fails if and only if G_d has triangles: in a half-one, two non-unital edges make the short sides in a degenerate triangle if and only if have the third side unital. Having G_d a tree certainly guarantees G_d triangle-free.

Proposition 4.1.1. *G_d a tree but not a path makes d extreme.*

Proof. G_d has a vertex v with degree at least 3. So (at least) three non-unital edges I_1, I_2, and I_3 share the vertex v. Then $L(I_1)$, $L(I_2)$, and $L(I_3)$ form a 3-cycle in $L(G_d) = \Gamma_d$.

Having $L(G_d)$ connected, Corollary (3.6.12) makes d extreme. □

More generally, write $\Delta(G)$ for the maximum degree of a vertex in a graph G. The same argument (and the observation above about co-products) gives:

Proposition 4.1.2 (Kehoe, 2018). *G_d triangle-free and $\Delta(C) \geq 3$ for every component C in G_d makes d extreme.* □

We'll use the result about trees to calculate a lower bound on the number of extreme half-one metrics. At the same time it will allow us to compare, for a given n, the cardinality of all half-one metrics and all metrics derived from partitions. Write $B_n = |\Pi_n|$ for the Bell numbers.

Proposition 4.1.3 (Kehoe, 2018). *$|ex(\mathcal{H}_n)| > n^{n-2} - \frac{n!}{2}$ for all n and $|ex(\mathcal{H}_n)| > B_n$ for n sufficiently large*

Proof. By Cayley's formula, n points admit n^{n-2} spanning trees; subtracting off the spanning paths, for $n \geq 2$,

$$n^{n-2} - \frac{n!}{2} < |\mathrm{ex}(\mathcal{H}_n)|\,.$$

Berend and Tassa [3] given an upper bound for the Bell numbers B_n:

$$B_n < \left(\frac{.792n}{\ln(n+1)}\right)^n.$$

Thus,

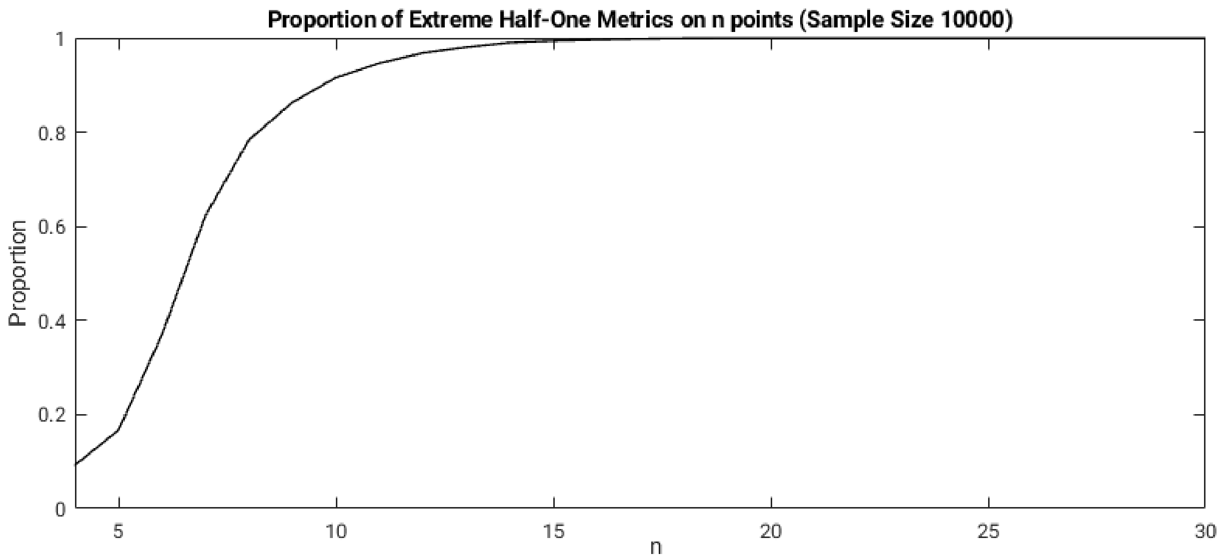
$$\frac{B_n}{|\mathrm{ex}(\mathcal{H}_n)|} < \frac{\left(\frac{.792}{\ln(n+1)}\right)^n n^2}{1 - \frac{n!}{2n^{n-2}}} \to 0$$

So that eventually,

$$B_n < |\mathrm{ex}(\mathcal{H}_n)|\,.$$

□

Exhaustive enumereation shows that the number of extreme half-one metrics first exceeds the number of partitions at $n = 5$. We can go further. Computer experiments strongly suggest the extremality of almost every half-one metric on n points for n large.



Conjecture 4.1.4. $|\mathcal{H}_n| \sim 2^{\binom{n}{2}}$ *i.e.* $\lim_{n\to\infty} \frac{|\mathcal{H}_n|}{2^{\binom{n}{2}}} = 1$.

Heuristic: Let $d \in \mathcal{H}_n$ and pick a half-length edge $[i,j] \in \mathcal{N}_d$. Now, pick another two distinct points l, k that are not endpoints of $[i,j]$. If we pick edge-lengths from the set $\left\{\frac{1}{2}, 1\right\}$ with equal probability for the remaining edge lengths on the complete graph with vertices $\{i,j,k,l\}$ with probability $\frac{1}{16}$ we obtain the extreme mid-point metric m_4. Thus we can expect $[i,j]$ to lie in $\frac{1}{16} \cdot \binom{n-2}{2}$ extreme midpoint metrics. n large then makes it extremely likely that every non-unital edge of d lies in an extreme mid-point metric, making d itself extreme.

4.2 The Upper Half of $\bar{\mathcal{M}}_n$

By the upper half of $\bar{\mathcal{M}}_n$ we mean the convex body $\bar{\mathcal{M}}_n^{\geq \frac{1}{2}} = \bar{\mathcal{M}}_n \cap [\frac{1}{2}, 1]^n$. Metrics living in this section of the body have a particularly nice decomposition.

Theorem 4.2.1 (Kehoe, 2018)**.** *Let $d \in \bar{\mathcal{M}}_n^{\geq \frac{1}{2}}$ then d arises as a convex combination of extreme $2^i 3^j$-den metrics where $0 \leq i + j \leq \lfloor \frac{m+1}{2} \rfloor$*

We begin with a result about decomposing mere vectors.

Lemma 4.2.2. *Let $p \in \left[\frac{1}{2}, 1\right]^n$ with $p_1 \leq \cdots \leq p_n$. Then we can express p as a convex combination of $n+1$ vectors h_i*

$$h_{ik} = \begin{cases} \frac{1}{2} & 1 \leq k < i \\ 1 & i \leq j \leq n \end{cases}$$

using coefficients ξ_k given by,

$$\xi_k = \begin{cases} 2\left(p_1 - \frac{1}{2}\right), & k = 1 \\ 2\left(p_k - p_{k-1}\right), & 2 \leq k \leq n \\ 2\left(1 - p_n\right), & k = n+1 \end{cases}$$

Proof. Basically the lemma just says that the points in a simplex arise as convex combinations of the extreme points, but we have use for the precise numerics.

We'll solve the system,

$$p = \sum_{i=1}^{n+1} \xi_i h_i$$
$$1 = \sum_{i=1}^{n+1} \xi_i$$

where $\xi_i \geq 0$. This yields the augmented matrix

$$\left(\begin{array}{cccc|c} 1 & 1 & \cdots & 1 & 1 \\ 1 & \frac{1}{2} & \cdots & \frac{1}{2} & p_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 1 & \cdots & 1 & \frac{1}{2} & p_n \end{array}\right)$$

By subtracting row $i-1$ from row i for $3 \leq i \leq n+1$, subtracting half of row 1 from row 2, and then subtracting row $n+1$ from row 1 we obtain the reduced augmented matrix

$$\left(\begin{array}{cccc|c} \frac{1}{2} & 0 & \cdots & 0 & p_1 - \frac{1}{2} \\ 0 & \frac{1}{2} & \cdots & 0 & p_2 - p_1 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{1}{2} & 1 - p_n \end{array}\right)$$

We then solve for the coefficients ξ_i as

$$\xi_i = \begin{cases} 2\left(p_1 - \frac{1}{2}\right) & , i = 1 \\ 2\left(p_i - p_{i-1}\right) & , 2 \leq i \leq n \\ 2\left(1 - p_n\right) & , i = n+1 \end{cases}$$

The non-negativity of these coefficients follows from $p_i \geq p_{i-1}$ and $p \in \left[\frac{1}{2}, 1\right]^n$.

□

Of course, now by permutating, we can express any vector $p \in \left[\frac{1}{2}, 1\right]^n$ as a convex combination of half-one vectors.

Applying this result to $\bar{\mathcal{M}}_n$ we obtain the corollary,

Corollary 4.2.3. $\bar{\mathcal{M}}_n^{\geq \frac{1}{2}} = \left[\frac{1}{2}, 1\right]^m$

Proof. Every half-one vector $h \in \mathbb{R}^m$ is a half-one metric on n points, and a convex combination of metrics is again a metric. Or we could simply observe that for any $d \in \left[\frac{1}{2}, 1\right]^m$, $d_{ij} + d_{jk} \geq \frac{1}{2} + \frac{1}{2} = 1 \geq d_{ik}$ so d automatically satisfies the triangle inequality. □

The rest of the story comes down to decomposing the *non-extreme* half-one metrics.

Proposition 4.2.4 (Kehoe, 2018)**.** *Let d be a non-extreme half-one metric with Γ_d connected. Then d equals the average of either two partitions, a partition and extreme positive-definite 3-den, or two extreme positive-definite 3-dens.*

Proof. Let (τ, E_0) be a rooted spanning tree for Γ_d and define $l(E)$ as the tree distance from the root E_0 to E in τ. By Proposition (3.7.2) every perturbation of d has the form $p(\varepsilon) = \varepsilon\eta$ with η the sign function,

$$\eta(E) = \begin{cases} (-1)^{l(E)} & , E \in \mathcal{N}_d \\ 0 & , E \in \mathcal{U}_d \end{cases}$$

Set $d_\varepsilon := d + p(\varepsilon)$. We now find the interval I over which d_ε is a metric. We have five non-trivial determining cases to consider for checking the triangle inequality.

1. $\frac{1}{2} + \varepsilon \leq \left(\frac{1}{2} + \varepsilon\right) + \left(\frac{1}{2} + \varepsilon\right) \implies -\frac{1}{2} \leq \varepsilon$

2. $\frac{1}{2} + \varepsilon \leq \left(\frac{1}{2} - \varepsilon\right) + \left(\frac{1}{2} + \varepsilon\right) \implies \varepsilon \leq \frac{1}{2}$

3. $\frac{1}{2} - \varepsilon \leq \left(\frac{1}{2} + \varepsilon\right) + \left(\frac{1}{2} + \varepsilon\right) \implies -\frac{1}{6} \leq \varepsilon$

4. $\frac{1}{2} + \varepsilon \leq 1 + 1 \implies \varepsilon \leq \frac{3}{2}$

5. $1 \leq 1 + \left(\frac{1}{2} + \varepsilon\right) \implies \varepsilon \leq \frac{1}{2}$

In general we get a metric d_ε for $-\frac{1}{6} \leq \varepsilon \leq \frac{1}{6}$.

Should we have every equilateral triangle of d uniformly signed (case 1 or opposite), we get a metric d_ε even for $-\frac{1}{2} \leq \varepsilon \leq \frac{1}{2}$. Letting $\varepsilon \in \left\{-\frac{1}{2}, \frac{1}{2}\right\}$, we obtain two partitions which average to d.

Suppose now that an equilateral triangle $[i, j, k]$ falls into case 2. Then by letting $\varepsilon = -1/6$ we obtain a 3-den metric $\tilde{d}$ such that,

$$\tilde{d}_{ij} = \tilde{d}_{jk} = \tfrac{1}{3} \quad , \quad \tilde{d}_{ik} = \tfrac{2}{3}$$

Since $\dim(P(d)) = 1$, d must lie on an edge of $\bar{\mathcal{M}}_n$ and thus making $\tilde{d}$ extreme. The proposition follows by similar arguments for the other cases. □

Theorem 4.2.5 (Kehoe, 2018)**.** *Let d be a non-extreme half-one metric such that Γ_d has exactly $N \geq 1$ non-trivial components and no isolated points. Then d arises as a convex combination of $2N$ extreme $2^i 3^j$-den metrics for $0 \leq i + j \leq N$. In the case that Γ_d contains isolated points d will be a convex combination of $2(N+1)$ such metrics.*

Proof. Suppose Γ_d has multiple components and let $Z_d \subset \Gamma_d$ be the set isolated points of Γ_d. Given an edge $E \in Z_d$ we can rigidify E by deforming its distance to 0 or 1. Isolated, E lies only in triangles where the adjacent edges either both have length 1/2 or both 1. Hence, the deformation preserves metricity. Deforming an isolated point to either 0 or 1 will have the effect of collapsing or joining (as nodes in the graph) the adjacent edges in every triangle in which the isolated point lies. Of course the original metric equals the average of these two deformations. So without loss of generality we can assume that Γ_d contains no isolated points.

Decompose Γ_d into non-rigid and rigid components as,

$$\Gamma_d = \kappa^1_{NR} \sqcup \cdots \sqcup \kappa^D_{NR} \sqcup \kappa^1_R \sqcup \cdots \sqcup \kappa^{D'}_R$$

where "NR" and "R" denote non-rigid and rigid respectively. Observe that we may view d as the co-product of metrics associated to these components. Moreover decomposition of one "summand" in the co-product yields a decomposition of the metric as a whole, just by leaving all distances outside the component fixed. Thus we may apply (4.2.4) one component at a time.

Define the perturbed metric $d_\varepsilon = d + p(\varepsilon)$ where,

$$p(\varepsilon) = \begin{cases} \varepsilon\eta_i(E), & E \in \kappa^i_{NR} \\ 0, & \text{otherwise} \end{cases}$$

with η_i defined as in Corollary (3.7.4). Let α and β be the minimum and maximum value respectively that ε can obtain so that d_ε enjoys metricity; compactness of $\bar{\mathcal{M}}_n$ guarantees their existence. If d lives in an r-face of $\bar{\mathcal{M}}_n$, we d arises as a convex combination d_α and d_β both living in a r'-face of $\bar{\mathcal{M}}_n$ where $r' < r$. We then apply the same process to both d_α and d_β, continuing until we obtain a set S of extreme metrics. Since each metric in the process equals a convex combination of the two following, d will live in the convex hull of S.

The process just described takes no more than D steps, the number of non-rigid components of Γ_d. In the case that d has isolated points, this only adds one more step while not changing the denominators involved. Thus we bound the number of denominator altering steps by the number D' of non-rigid components with at least 2 points.

We can now calculate the type of denominators we could encounter in the prescribed process. At any give step, we've focused on one non-rigid component to which we can apply (4.2.4). Write $[i, j, k]$ for a non-degenerate triangle in that component that becomes degenerate upon perturbation. Write $\Delta = d_{ij} + d_{jk} - d_{ik} > 0$. By (4.2.4)

$$\Delta - q\varepsilon = 0 \qquad \text{where} \qquad q = \pm 1, \pm 2, \text{ or } \pm 3$$

so that,

$$\varepsilon = \frac{\Delta}{q} \qquad q = \pm 1, \pm 2, \text{ or } \pm 3.$$

Hence if d is an r-den metric then the perturbed metric is a r'-den for r' a divisor of rq.

Hence d is a convex combination of 2^i3^j-den metrics where $0 \leq i+j \leq D'$. □

We now come to the proof of the main theorem,

Theorem (Kehoe, 2018)**.** *Let $d \in \bar{\mathcal{M}}_n^{\geq \frac{1}{2}}$ then d is a convex combination of extreme 2^i3^j-den metrics where $0 \leq i+j \leq \lfloor \frac{m+1}{2} \rfloor$*

Proof. Let $d \in \bar{\mathcal{M}}_n^{\geq \frac{1}{2}}$ then by Corollary (4.2.3) d arises as a convex combination of half-one metrics. By Corollary (4.2.5) every non-extreme half-one metric has a decomposition as a convex combination of extreme 2^i3^j-den metrics. We can give an upper bound on $i+j$ by considering a partition of m with the maximum number of parts and at least one isolated point. Without loss of generality, we can assume m odd, so that we pair off edges leaving one isolated. Thus, we have that $i+j \leq \frac{m+1}{2}$. This completes the proof.

□

4.3 Neighbors of Half-One Metrics

Given an extreme metric d we would like to find its neighbors in $\bar{\mathcal{M}}_n$, meaning extreme points of $\bar{\mathcal{M}}_n$ connected to d by an edge of the metric body. Neighbors of d arise by choosing $\binom{n}{2} - 1$ linearly independent active constraints from (A_d, a_d). Such a choice will generate a 1-dimensional affine subspace L of $\mathbb{R}^m$. If L intersects $\bar{\mathcal{M}}_n$ at more than d, then L will lie on an edge of d and hence generate a neighbor of d.

If we choose d to be a half-one metric and decide to keep all of the unital constraints, then an appropriate choice of constraints for a neighbor of d will be equivalent to picking a spanning tree for a component in the edge graph of d. Of course, some spanning trees will not generate any neighbor due to infeasibility.

For starters, assume Γ_d connected. d a half-one metric makes Γ_d short-sided. Now pick a spanning tree τ for Γ_d. From previous results we have $\{e_T\}_{T\in\tau}$ linearly independent, in particular over V_τ. We can generate neighbors of d by keeping the same unital edges as d and perturbing non-unital edges according to the perturbation $p_{(\tau,E_0)} : \mathbb{R} \to \mathbb{R}^m$ associated to the sign function $\eta_{(\tau,E_0)}$ for (τ, E_0) with E_0 some arbitrarily chosen root. Thus we perturb by $p(\varepsilon) = \varepsilon\eta$ where

$$\eta_{(\tau,E_0)}(E) = \begin{cases} (-1)^{l(E)}, & E \in \mathcal{N}_d \\ 0, & E \in \mathcal{U}_d \end{cases}$$

and l is the associated tree distance from the root E_0. Note that if one changes the root of τ we obtain the same sign funtion η and hence the same perturbation p up to sign. Indeed,

Proposition 4.3.1. *Let (τ, E_0) and (τ, E_0') be the same tree with different roots, then $\eta_{(\tau,E_0)} = \pm\eta_{(\tau,E_0')}$.*

Proof. 2-color the nodes of τ (so that adjacent nodes don't receive the same color). The even parity of the length between nodes means precisely that the nodes have the same color. Thus $\eta_{(\tau,E_0)}$ and $\eta_{(\tau,E_0')}$ will differ by a sign or not according to whether or not E_0 and E_0' receive the same color. □

To each tree spanning tree τ we have an associated 1-dimensional affine subspace L_τ through d. We want to know when L_τ is the affine span of an edge of $\bar{\mathcal{M}}_n$, i.e. when τ generates an extreme neighbor.

Definition 4.3.2. Fix an edge $T = [i, j, k] \in E(\Gamma_d)$ in the edge graph, so a degenerate triangle. Call T *positively signed* with respect to rooted spanning tree (τ, E_0) if $\eta_{\tau,E_0}([i, j]) = \eta_{\tau,E_0}([j, k]) = +1$; call T *negatively signed* if $\eta_{\tau,E_0}([i, j]) = \eta_{\tau,E_0}([j, k]) = -1$. Otherwise call T *polar.*

Theorem 4.3.3 (Kehoe, 2019)**.** *Let $d \in ex(\mathcal{H}_n)$ with Γ_d connected. Given a spanning tree $\tau \subset \Gamma_d$ signed by any given root, τ generates an extreme neighbor d_τ of d if and only if $\tau^C = \Gamma_d \setminus \tau$ does not contain both a positively signed edge and a negatively signed edge. If τ does generates an extreme neighbor, d_τ take the form of a partition metric or* 3*-den true metric.*

Proof. Suppose τ^C contains a positively signed edge T_+ and a negatively signed edge T_-. Let p_τ denote the perturbation associated to τ for some arbitrarily chosen root. Now we have two degenerate triangle for d, $T_+ = [i,j,k]$ and $T_- = [u,v,w]$ and the perturbation $p_\tau(\varepsilon)$ violates feasibility for $d_{ij} + d_{jk} \geq 1$ if $\varepsilon < 0$ and for $d_{uv} + d_{vw} \geq 1$ if $\varepsilon > 0$. That leaves only $\varepsilon = 0$ so that L_τ does not generate an extreme neighbor.

If τ^C does not contain both a positively signed edge and a negatively signed edge, without loss of generality assume that τ^C contains only positively signed edges and polar edges. If $\varepsilon < 0$ none of the triangle inequalities associated to positively signed edges of τ^C get violated, and thus the perturbation p_τ will generate a ray which intersects $\mathcal{M}_n$ on an edge $[d, d_\tau]$. As per the proof of Theorem (4.2.4), d_τ must take the form either of partition metric or 3-den true metric. □

If d does not have a connected edge graph, we still obtain a 1-dimensional affine space in $\mathbb{R}^m$ through d via a spanning tree τ for a single component κ of Γ_d. Since all the components of d are rigid, we "loosen" up a component by removing edges until we have our τ.

Corollary 4.3.4 (Kehoe, 2019)**.** *Suppose $d \in ex(\mathcal{H}_n)$ with Γ_d disconnected. Let τ be a spanning tree for a component κ of Γ_d. Then τ generates an extreme neighbor d_τ of d if and only if $\kappa \setminus \tau$ does not contain both a positively signed edge and a negatively signed edge. In the case that τ generates an extreme neighbor, we have that d_τ is either a cover of an extreme half-one metric or an extreme 6-den metric.*

Proof. Viewing the metric as a co-product, we apply the previous theorem to one summand. In the case that d_τ exists the edge lengths associated to the component κ will either be 0's and 1's or $\frac{1}{3}$'s and $\frac{2}{3}$'s. By taking a common denominator with the other half length edges we obtain either denominator 2 or 6 respectively. □

If we decide to keep all but one unital constraint, say $d_{I_0} = 1$, we get a few possibilities for choosing $\binom{n}{2} - 1$ active constraints of (A_d, a_d) for generating possible neighbors.

1. We could choose a spanning colony for Γ_d.

2. We could choose germs for all but one component κ, and for κ choose any spanning subgraph with an even cycle.

3. We could choose germs for all but one component κ, and for κ we choose any spanning tree τ. Finally, for some germ g_i contained in component κ_i we add an edge of $\kappa_i \setminus g_i$ to g_i to obtain a subgraph λ of κ_i.

Determining the denominators of these types neighbors remains to do. With Proposition (3.6.6) in mind, we would need to develop machinery to keep track of the unital edges associated to edges in the edge graph.

CHAPTER 5

Experimental Results

This chapter reports on results from some ongoing experiments concerning the geometry of the metric body $\bar{\mathcal{M}}_n$.

5.1 Extreme Points

Computation can identifying extreme points in $\bar{\mathcal{M}}_n$. We calculated the following lists of extreme points of $\bar{\mathcal{M}}_n$ using Multi-parametric Tool Box (mpt3), a MATLAB based computational geometry software. We enumerated the vertices of $\bar{\mathcal{M}}_n$ using the simplex method, see [5]. Here we list the positive-definite extreme metrics, scaled to integer points, as row vectors in $\mathbb{R}^m$. In addition, we mod out by the appropriate symmetric group to remove equivalent extreme points on a set of the same cardinality.

Scaled extreme true metrics on 4 points modulo S_4.

$$|\mathrm{ex}\left(\bar{\mathcal{M}}_4\right)| = 19$$

Table 4.1.1

2	1	2	1	2	1
1	1	1	1	1	1

Scaled extreme true metrics on 5 points modulo S_5.

$$|\mathrm{ex}\left(\bar{\mathcal{M}}_5\right)| = 259$$

Table 4.1.2

1	2	1	1	1	2	2	1	1	2
1	2	1	1	1	2	2	1	2	2
1	2	1	1	2	2	2	2	1	2
1	2	1	1	2	2	2	2	2	2
1	2	1	2	2	2	1	1	1	2
1	2	2	2	1	1	1	2	2	2
1	1	1	1	1	1	1	1	1	1

Scaled extreme true metrics on 6 points modulo S_6. $|\text{ex}\left(\bar{\mathcal{M}}_6\right)| = 27263$

Table 4.1.3

1	2	1	1	1	1	2	2	2	2	1	1	1	2	2
1	2	1	1	1	1	2	2	2	2	1	2	2	2	2
1	2	1	1	1	1	2	2	2	2	1	1	2	2	2
1	2	1	1	1	1	2	2	2	1	1	1	2	2	2
1	2	1	1	2	1	2	2	3	1	1	2	2	1	1
1	3	1	3	3	2	2	4	4	4	4	2	2	4	2
1	2	1	2	1	1	2	3	2	3	2	1	3	2	3
1	3	1	3	3	2	2	4	4	4	2	2	4	4	4
1	3	1	3	3	2	2	4	4	4	2	2	2	4	4
2	4	1	2	2	2	3	4	4	3	2	4	3	1	4
1	3	1	2	2	2	2	3	3	2	1	3	3	1	3

1	3	1	2	2	2	2	3	3	3	1	2	3	1	3
1	3	1	2	2	2	2	3	3	2	1	3	3	1	2
1	3	1	2	2	2	2	3	3	3	1	3	3	1	2
1	2	1	1	1	1	2	2	2	1	2	2	1	2	2
1	2	1	1	1	1	2	2	2	2	1	2	1	2	1
1	2	1	1	1	1	2	2	2	1	1	2	2	1	1
1	3	1	2	2	2	2	3	3	2	1	3	1	3	2
1	2	1	1	1	1	2	2	2	1	1	2	1	2	1
1	2	1	1	1	1	2	2	2	2	2	2	2	2	2
1	2	1	1	1	2	2	2	2	2	2	2	2	2	2
1	1	1	1	1	1	2	2	2	1	1	2	2	2	2
1	2	1	1	2	1	2	2	2	1	1	2	2	2	1
1	2	1	1	2	1	2	2	2	1	2	1	2	2	1
1	2	1	1	2	1	2	2	2	1	1	2	2	2	2
1	2	1	1	2	1	2	2	2	2	1	2	2	1	1
1	2	1	1	2	1	2	2	2	1	1	2	2	1	1
1	2	1	1	2	1	2	2	2	1	2	2	2	1	2
1	2	1	1	2	1	2	2	2	2	1	1	2	2	2
1	2	1	1	2	1	2	2	2	2	1	1	1	1	2
1	2	1	1	2	1	2	2	2	1	1	2	1	1	2
2	4	2	2	1	4	4	4	1	2	4	3	4	3	3
2	4	2	2	1	4	4	4	1	4	4	3	4	3	3
1	3	2	1	2	2	3	2	1	1	3	3	3	3	3
1	2	2	1	3	1	3	2	2	3	3	3	1	3	2
1	2	1	1	2	1	2	2	1	1	1	2	2	1	1

1	2	1	1	1	2	2	2	2	1	2	2	1	2	2
1	3	1	2	2	2	2	3	3	2	3	3	1	3	2
1	3	1	2	2	2	2	3	3	2	3	3	1	3	3
1	3	2	1	2	3	3	2	3	3	2	3	1	3	3
1	1	1	1	1	1	2	2	2	2	2	1	2	1	1
1	3	2	1	2	3	3	2	3	3	3	1	1	2	3
1	3	2	1	2	3	3	2	3	2	3	1	1	3	3
1	2	1	1	1	2	2	2	2	2	1	2	1	2	1
1	4	1	3	3	3	2	4	4	4	2	4	2	4	2
1	4	1	3	3	4	2	4	4	4	2	4	2	4	2
1	2	1	2	1	1	2	3	2	3	3	1	3	2	3
1	2	2	2	2	1	3	3	3	4	2	2	4	4	4
1	2	1	1	1	1	2	2	2	1	2	2	1	2	1
1	2	1	1	2	2	2	2	1	2	1	1	2	2	2
1	1	1	1	2	1	2	2	1	1	1	1	2	1	2
1	2	1	1	2	2	2	2	1	1	2	1	2	1	2
1	2	1	1	2	2	2	2	1	1	1	1	1	1	2
1	2	1	1	2	2	2	2	1	2	1	2	2	1	2
1	2	1	1	2	2	2	2	1	2	2	2	2	1	2
1	2	1	1	2	1	2	2	1	2	2	2	2	2	2
1	2	1	1	2	2	2	2	2	1	2	2	1	1	1
1	2	1	1	1	2	2	2	1	2	2	1	1	2	1
1	2	1	1	2	2	2	2	2	2	1	2	2	2	2
1	2	1	1	2	2	2	2	2	2	1	2	2	1	2
1	2	1	1	2	2	2	2	2	2	1	1	2	2	2
1	2	1	1	2	2	2	2	2	2	2	2	2	2	2
1	2	1	1	1	2	2	2	1	2	1	1	1	2	1

1	1	1	1	1	1	2	2	1	2	1	2	1	1	2
1	2	1	1	2	2	2	2	1	2	1	1	1	2	2
1	2	1	2	2	2	2	2	1	2	2	2	1	2	1
2	1	2	2	2	1	2	2	2	1	1	1	2	2	2
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

For $n > 6$ we found the number of extreme points too large to compute in a reasonable amount of time; the output grows at least exponentially (see the previous section) and the simplex algorithm takes longer since the size of the instances grows quadratically.

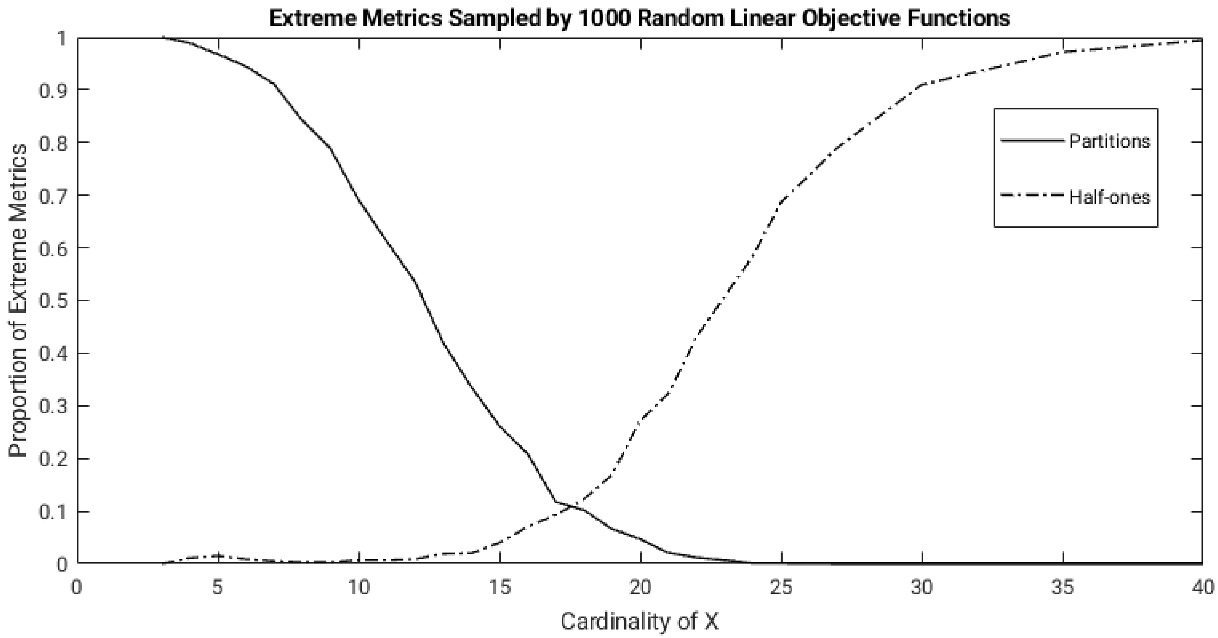
5.2 Half-ones Optimize Linear Objectives

Given a linear form $f \in \mathbb{R}^m$ with $m = \binom{n}{2}$ we can ask at what points of $\bar{\mathcal{M}}_n$ does f obtain its maximum. Compactness of $\bar{\mathcal{M}}_n$ guarantees that f attains its maximum. In particular, the maximum principle in convexity theory says that a convex function f attains its maximum, d_*, on the boundary of $\bar{\mathcal{M}}_n$. So d_* lies in a face of $\bar{\mathcal{M}}_n$. Reapplying the maximum principle and restricting f to ever lower dimensional subfaces, we eventually obtain an extreme metric.

Experiments suggest that d_* almost always equals an extreme half-one metric for large n (figure below). Earlier we saw that half-one metrics outnumber partitions (see Theorem (4.1.3)). These experimental results lead one to wonder if half-one metrics simply outnumber all q-den metrics for $q > 2$? We investigate this question.

Call a positive rational point $d \in [0, 1]^m$ with denominator q a *q-level point*, and denote the set thereof by I^n_q. Certainly $|I^n_q|$ increase with q. Rational points $x \in [0, 1]^m$ that represent metrics must satisfy $3\binom{n}{3}$ triangle inequalities. The ratio of triangle inequalities to edges for metrics on n points equals $n-2$. That makes it increasingly difficult as q grows for a random rational point x to come out a metric. Of course 2-level points represent metrics, the half-ones, automatically.

To interrogate the probability that a random q-level point comes out a metric, we can calculate the probability that a triple in a given q-level point satisfies the triangle inequalities.



Consider triples of numbers from the set $\{1,\dots,q\}$ where $q \in \mathbb{Z}^+$. Call (a_1, a_2, a_3) feasible if $a_i + a_j \geq a_k$ for all distinct i, j, k. We count feasible triples by counting their complement. Fixing $a_3 = k \leq q$, consider bi-partitions, meaning equations $a_1 + a_2 = s$, of $s < k$. s has exactly s bi-partitions, so that the number of sums where $a_1 + a_2 < k$ equals

$$\begin{aligned}\sum_{s=1}^{k-1} s &= \frac{k(k-1)}{2} \\ &= \binom{k}{2}\end{aligned}$$

Then summing k from 1 to q yields the number of triples that fail one-triangle inequality

$$\begin{aligned}\sum_{k=1}^{q} \frac{k(k-1)}{2} &= \frac{q(q+1)(q-1)}{6} \\ &= \binom{q+1}{3}.\end{aligned}$$

A triple that fails one triangle inequality automatically satisfies the other two. Thus the number of infeasible triples equals

$$\frac{q(q+1)(q-1)}{2}$$

making the number of feasible triples

$$q^3 - \frac{q(q+1)(q-1)}{2} = \frac{1}{2}q\left(q^2+1\right)$$

Now we get the probability of feasibility for a random triple as

$$\begin{aligned} P_q &= 1 - \frac{1}{2}q\left(q^2+1\right) \\ &= \frac{1}{2} - \frac{1}{2q^2} \end{aligned} \tag{5.2.1}$$

Letting $q \to \infty$ we obtain the asymptotic probability for (rational) triple feasibility,

$$P_{\mathbb{Q}} = \frac{1}{2}\ .$$

From a geometric viewpoint, we just calculated the volume of $\bar{\mathcal{M}}_3$. Indeed, if we pick all the rational points in the unit cube with denominator q then ask for the proportion of points that land in $\bar{\mathcal{M}}_3$ we obtain an approximate volume. Taking a limit as $q \to \infty$ gives the exact value. [2]. If $x < y < z$, we have (x, y, z) feasible if and only if we have $(z-x, z-y, z)$ infeasible. This involution thus yields an alternative derivation of $P_{\mathbb{Q}} = \frac{1}{2}$.

We are almost ready to give a bound on the probability for a q-level point to be a metric. First we state an important theorem on the independence of triangles.

Theorem (Spencer 1968)**.** *A set of n points can be covered by $C(n)$ independent (edge-*

disjoint) triangles where

$$C(n) = \begin{cases} \frac{1}{3}\binom{n}{2} & , \quad n \not\equiv 5 \mod 6 \\ \frac{1}{3}\binom{n}{2} - 1 & , \quad n \equiv 5 \mod 6 \end{cases}$$

and $C(n)$ is the maximum possible value.

Fix a maximal independent triangle covering $\{T_1, \ldots, T_{C(n)}\}$ of X and let P denote the standard probability measure on $[0,1]^m$ (Lebesgue measure). Pairing Spencer's theorem with (5.2.1),

$$\begin{aligned} P\left(I_q^m \cap \mathcal{M}_n\right) &\leq P\left(\left\{d \in I_q^m \,|\, T_j \text{ is feasible for all } j\right\}\right) \\ &= \left(\frac{1}{2} - \frac{1}{2q^2}\right)^{C(n)} \end{aligned}$$

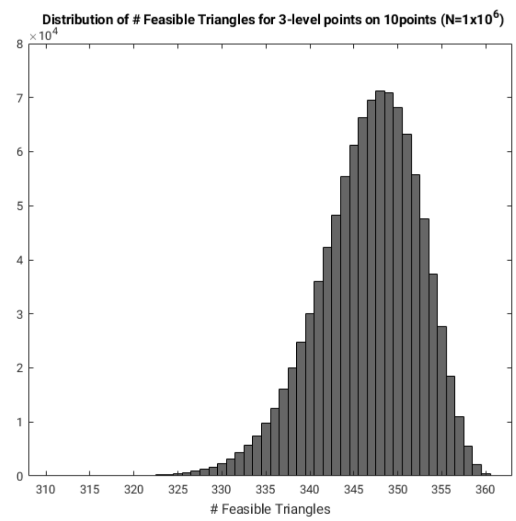
From this we can see that, as n grows large, it becomes increasingly unlikely for a random q-level point to come out a metric. In fact we can use this bound to prove the following

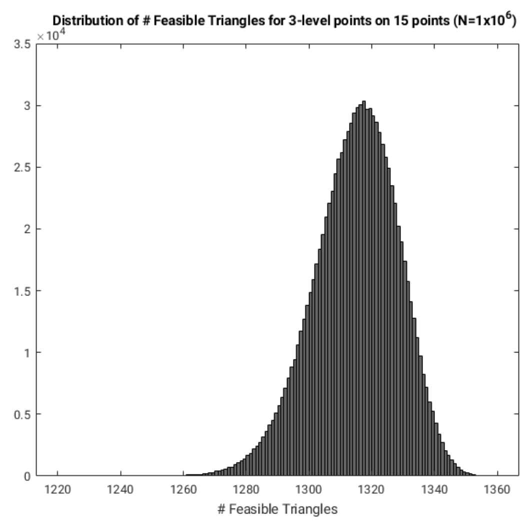
Corollary 5.2.2. *$|I_q^n \cap \mathcal{M}_n| \leq 2\left(\sqrt[3]{\frac{q^3}{2} - \frac{q}{2}}\right)^{\binom{n}{2}}$ and $Vol(\mathcal{M}_n) \leq 2\left(\frac{1}{\sqrt[3]{2}}\right)^{\binom{n}{2}}$. In particular $Vol(\mathcal{M}_n) \to 0$ as $n \to \infty$.*

Proof. $|I_q^n \cap \mathcal{M}_n| = P\left(I_q^n \cap \mathcal{M}_n\right) q^{\binom{n}{2}}$. For the second part, take the limit as $q \to \infty$ and then the limit as $n \to \infty$. □

However unlikely a q-level point to come out a metric, q-level points generally far outnumber half-one metrics. The bound above doesn't have the strength to make metrics arising from q-level points rare in $\bar{\mathcal{M}}_n$ compared to half-one metrics. The bound does make trying to sample extreme q-level metrics via a uniform distribution on q-level points is nearly useless when n is large. We need a different strategy for obtaining an approximation of the number of extreme q-level metrics.

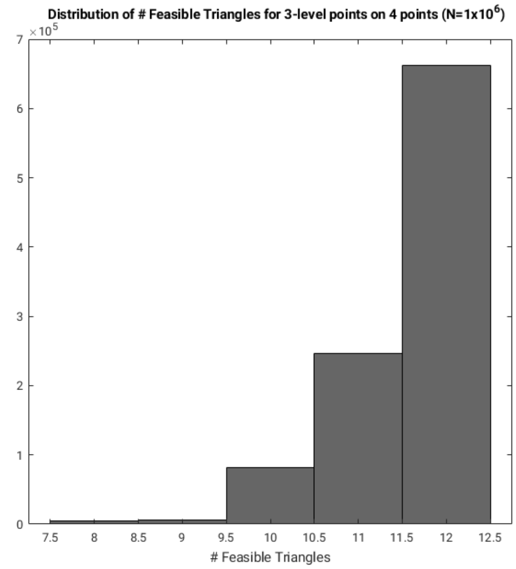
Idea: As n grows large, two triangles chosen at random become increasingly unlike to share an edge. Sharing no edges makes their feasibility or not in q-level point independent. By the central limit theorem the distribution on the number of feasible triangles and degenerate triangles becomes approximately normal. (We show an example of the progression of this normalization for a sample of uniformly chosen 3-level points on the next page). Why not use these approximating Gaussian distributions to make estimates on the number of extreme metrics?

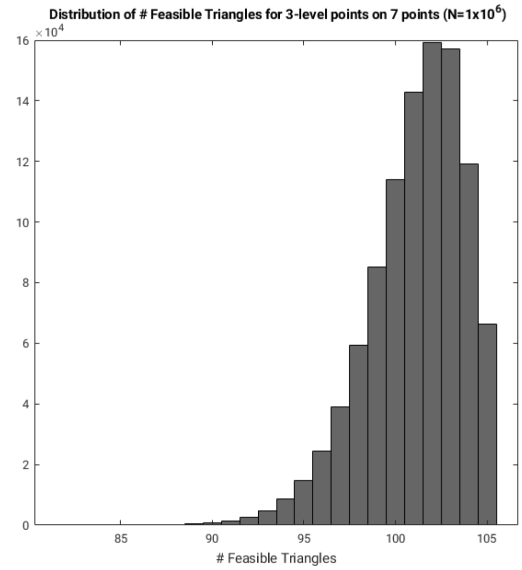




To make such estimates, we'll need an indicator of extremality. Certainly we can check the rank of the active matrix for each q-level point, but making many rank calculations for large matrices costs time. Instead we use the number of degenerate triangles for a q-level metric as an indicator of extremality. Indeed, for q-level point to come out a basic solution requires at least $\binom{n}{2}$ active constraints. If a q-level metric has many degenerate triangles compared to the number of edges, extremality becomes likely.

Thus we want to search for those points both highly feasible and highly degenerate. A degenerate triangle in a q-level point makes the other two associated triangle inequalities automatically satisfied. Hence, at least for a triangle, degeneracy implies feasibility.





Question: Does degeneracy imply feasibility in general? What is the relationship between the two?

We answer this question with certainty in an extreme case.

Proposition 5.2.3. *For $d \in \mathbb{R}^{\binom{n}{2}}$ strictly positive, d having more than $\binom{n}{3}$ degenerate triangles makes d feasible.*

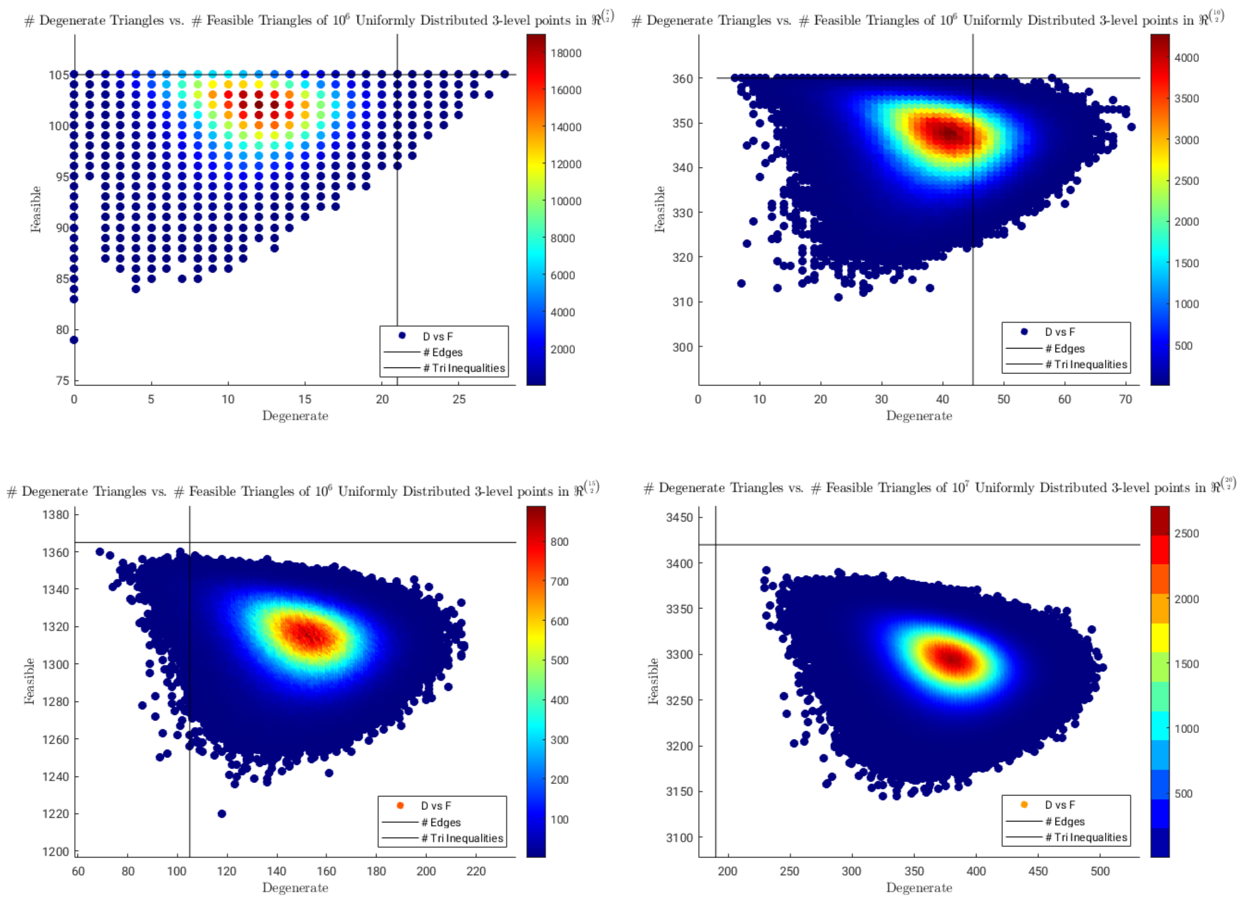
Proof. For strictly positive d, no triangle has more than one degenerate triangle inequality. By the pigeonhole principle, the condition makes every triangle of d degenerate and hence feasible. □

To investigate further we run an experiment with 3-level points to look for a correlation between feasibility and degeneracy.

Experiment 1: Pick N uniformly distributed q-level points, calculate their number of feasible triangles and degenerate triangles. Plot a frequency distribution, and search for high feasibility, high degeneracy points.

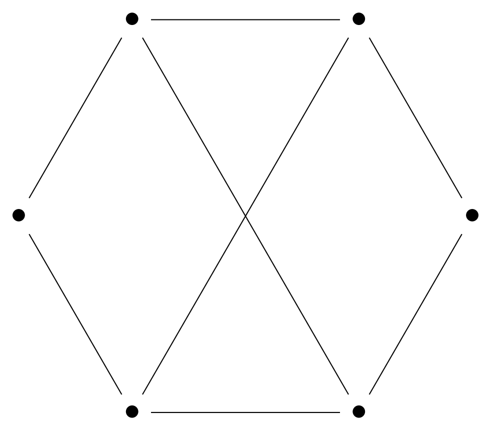
We plot the results of this experiment below for 3-level points in $\mathbb{R}^{\binom{n}{2}}$ for $n = 7, 10, 15, 20$

and $N = 10^6, 10^7$.



We can see that as n grows large degeneracy increases but at the cost of feasibility. What causes this drop in feasibility? Infeasibility for a 3-level triangle occurs only when the distances take the form $(\frac{1}{3}, \frac{1}{3}, 1)$. This infeasibility conflicts with degeneracies having the form $(\frac{1}{3}, \frac{2}{3}, 1)$. Each degeneracy of this type requires 1's and $\frac{1}{3}$'s, thereby increasing the probability that some other triangle comes out infeasible. In general, q-level points having an ample supply of degenerate triangles of the form $a + b = c$ where $a < b$ promotes the occurrence of infeasibility: if we encounter many distances of size a and c, we will likely come across a triangle where $a + a < c$.

We then have two ways in which we could obtain feasible 3-level points. Either limit the amount of $\frac{1}{3}$'s or limit the amount of 1's. Consider again the example of the Bow-tie:



The Bow-tie enjoys feasibility, has few unital edges (indeed just one), but many degenerate triangles. It serves as a prime example for generalization. One very important feature of the Bow-tie is that every triangle involving the unital edge is degenerate, this ensures feasibility of the Bow-tie and creates a "core" of degeneracies around the unital "stem" for the remaining lower level degeneracies $(\frac{1}{3}, \frac{1}{3}, \frac{2}{3})$ to link to. Define the set

$$\mathcal{B}^3_{n,1} = \left\{ d \in I^3_n \,|\, \exists! I \in \mathcal{E} : d_I = 1, \forall T \in \mathcal{T} : I \subset T \implies T \text{ is degenerate} \right\}$$

Since we disallow triangles of the form $(\frac{1}{3}, \frac{1}{3}, 1)$ every member of $\mathcal{B}^3_{n,1}$ defines a metric. Having degeneracies for every triangle formed off of the unital edge will create an anchor point for the $(\frac{1}{3}, \frac{1}{3}, \frac{2}{3})$ degenerate triangles, so that we will be able to emulate the nature of half-one metrics.

The discovery of the large class of extreme points given by the co-products of generalized bow-ties, indeed leads us to expect at least as many non-half-one extremes metrics as extreme half-one metrics. We expect not the particular abundance of half-one metrics, but something about the geometry of the body near a half-one metric that causes them to optimize randomly chosen linear objective functions so often. By this geometric feature we mean the sharpness of the extreme point, the size of the dual-cone at the extreme point. Of course this dual-cone size gives, up to normalization, the probability of the *particular* extreme point maximizing an random objective function. So we conjecture that the sum of the exterior solid polyhedral angles over all extreme half-one metrics on n points limits to the volume of an $\binom{n}{2}$-dimensional sphere of radius 1. For those familiar with the Gauss-Bonnet theorem, the convex polytope version of theorem replaces curvature on a manifold with the exterior solid polyhedral angles at vertices. So, from the optic of Gauss-Bonnet theorem, we mean to guess that extreme half-ones tend to optimize linear objection functions because they "eat" up almost all the curvature of the metric body. Random half-one tend to sit on far more hyperplanes that mere extremality requires, because they have so many degenerate triangles. We guess that intersecting so many "extra" half-spaces makes these points "sharp" and their dual cones large.

Calculating spherical polyhedral angles exactly and in general turn out a difficult task, but a bound on such angles may suffice when summing over a large family of vertices. We hope in the future to investigate the hyper-geometric series that come about from calculating such polyhedral angles.

CHAPTER 6

The Complex of Ultrametrics and Iterated Cycle Structures

In this chapter we shift focus to another important class of metrics, the ultrametrics, generated by the partition metrics. We will describe the structure of so called ultrametrics up to homotopy equivalence. This will require an investigation into the topology of the order complex of partition lattices. There, we will give a new proof of the Euler-characteristic of this complex, and extend this proof technique to the context of iterated cycle structures, an object analogous to chains in the lattice of partitions.

6.1 Definitions and Preliminaries

Definition 6.1.1. A *bounded-by-1 pseudoultrametric* means a function
$d : X \times X \to [0, 1]$ such that for any $x, y, z \in X$

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

2. $d(x, y) = d(y, x)$ (Symmetric)

3. $d(x, z) \leq \max\{d(x, y), d(y, z)\}$ (Strong Triangle Inequality)

The set of bounded-by-1 pseudoultrametrics will be denoted by $U(X)$ and simply U_n if $X \cong \{1, \ldots, n\}$. Using the same abbreviation as we did for bounded-by-1 pseudometrics, we'll call any member of U_n an ultrametric. A routine check shows that every ultrametric is in fact a metric. When looking for examples, partitions metrics form an important class of ultrametrics.

Satisfying the strong triangle inequality puts great a deal of restrictions on the number of degrees of freedom for choice of ultrametric. The proposition below makes this precise.

Proposition 6.1.2. *Every triangle $[x, y, z]$ for an ultrametric d is isosceles. Furthermore, if $d \in U_n$ then d has at most n distinct distances.*

Proof. Consider the triangle $[x, y, z]$ then we have the three simultaneous conditions,

$$d(x, z) \leq \max \{d(x, y), d(y, z)\}$$

$$d(x, y) \leq \max \{d(x, z), d(y, z)\}$$

$$d(y, z) \leq \max \{d(x, y), d(x, z)\}$$

We'll assume that we have at least two distinct distances, or else we would already have an isosceles triangle. Without loss of generality assume that $d(x, y) < d(x, z)$. By the third inequality $d(y, z) \leq d(x, z)$. If $d(y, z) = d(x, z)$)done. But if $d(y, z) < d(x, z)$ the first condition give $d(x, z) \leq d(x, y)$, contradiction.

We will prove the second part of the proposition by induction. The result hold trivially for $n = 1$ and $n = 2$ since triangles do not appear until $n = 3$; we'll take $n = 3$ as our induction base. Indeed, when $n = 3$ we have only one triangle, isosceles by the previous result. Hence we indeed have at most two distinct non-zero distances.

Now assume $d \in U_n$ and the proposition is true for $n - 1$. We can restrict d to the set $[n - 1] = \{1, \dots, n - 1\}$, so that by the induction hypothesis d has at most $n - 1$ distinct distances on $[n - 1]$. Consider now $1 \leq i \leq n$ and the necessarily isosceles triangles $[1, i, n]$.

Thus $d_{in} \in \{d_{1i}, d_{1n}\}$ or $d_{1i} = d_{1n}$.

In the first case, d_{in} either equals one of the $n - 1$ distances given by restriction or the single distance d_{1n}.

In the second case, d_{1n} equals one of the original $n - 1$ distances.

Assume w.l.o.g. that d_{in} defines a new distance. Consider the isosceles triangle $[i, j, n]$ where $1 \leq j \leq n - 1$. Then $d_{jn} \in \{d_{ij}, d_{in}\}$ lest d_{in} equal one of the original $n - 1$ distances.

That makes d_{jn} equal one of the n distinct distances already defined. In any case, d has at most n distinct distances, completing the induction. □

Given an ultrametric $d \in U_n$ and point $t \in I_d = [0,1] \setminus \text{Im}(d)$ called a *threshold*, we can define a new function d_t defined by,

$$d_t(i,j) = \begin{cases} 0, & d(i,j) < t \\ 1, & d(i,j) > t \end{cases}$$

Proposition 6.1.3 (Feldman, Kehoe, 2019)**.** *d_t is a partition metric.*

Proof. We'll prove that d_t defines an equivalence relation on $[n]$. Reflexivity and symmetry are immediate consequences of the definition of metric. Now suppose $x \sim y$ and $y \sim z$. Then $d(x,y) < t$ and $d(y,z) < t$ then by the strong triangle inequality,

$$\begin{aligned} d(x,z) &\leq \max\{d(x,y), d(y,z)\} \\ &< t. \end{aligned}$$

Hence $x \sim z$, proving transitivity. □

Now that we know ultrametrics give rise to partitions we would like to know the relationships between the partitions that arise from a given ultrametric by varying the threshold $t \in I_d$. Before we get into describing the relationships involved, we need to first learn about the structure of the lattice of partitions, a well-known object of study.

6.2 Shellable Posets and Their Order Complexes

In this section we develop the language and machinery to describe the structure of Π_n.

Definition 6.2.1. A *partially ordered set* (poset) $(\mathcal{P}, \leq)$ means a set $\mathcal{P}$ together with a relation $\leq$ satisfying the conditions below, for $a, b, c \in \mathcal{P}$.

1. $a \leq a$ (reflexivity);

2. $a \leq b$ and $b \leq a$ implies $a = b$ (anti-symmetry);

3. $a \leq b$ and $b \leq c$ implies $a \leq c$ (transitivity).

If context makes the relation $\leq$ clear, we call a poset by the name of the underlying set $\mathcal{P}$. In the case that $a \neq b$ we will say $a < b$. Note that we don't assume every pair of elements in a poset comparable.

We call a poset $\mathcal{P}$ *bounded* if it has a least element $\hat{0}$ and a greatest element $\hat{1}$. Given bounded $\mathcal{P}$ we define the sub-poset $\hat{\mathcal{P}} = \mathcal{P} \setminus \{\hat{0}, \hat{1}\}$.

Definition 6.2.2. A *chain* of length k in $\mathcal{P}$ between $a, b \in \mathcal{P}$ means a collection of elements $\{a_i\}_{i=1}^{k}$ such that $a = a_1 < \cdots < a_k = b$. We call a chain $\mathcal{C}$ between a and b *refinable* if we can find another chain $\mathcal{C}'$ between a and b such that $\mathcal{C} \subsetneq \mathcal{C}'$.

We call a finite poset *pure* if all maximal chains have the same length.

Proposition 6.2.3. *A pure poset $\mathcal{P}$ satisfies the* Jordan-Dedekind condition*: all unrefinable chains between two comparable elements a and b have the same length.*

Proof. Take two unrefinable chains $\mathcal{C}, \mathcal{C}'$ between a and b of length k and l respectively. Then $\mathcal{C}$ and $\mathcal{C}'$ extend to two maximal chains $\hat{\mathcal{C}}$ and $\hat{\mathcal{C}}'$ respectively. Without loss of generality we can choose chains $\hat{\mathcal{C}}$ and $\hat{\mathcal{C}}'$ so that

$$\hat{\mathcal{C}} \setminus \mathcal{C} = \hat{\mathcal{C}}' \setminus \mathcal{C}'.$$

Unrefinability guarantees that this set subtraction reduces $\hat{\mathcal{C}}$ and $\hat{\mathcal{C}}'$ by exactly the numbers of elements in $\mathcal{C}$ and $\mathcal{C}'$ respectively. If we let $n = \text{length}(\hat{\mathcal{C}})$, by purity we obtain the equation $n - l = n - k$ implying $l = k$. □

Definition 6.2.4. We call a finite, bounded, pure poset $\mathcal{P}$ *graded.* The *rank* $\rho(x)$ for $x \in \mathcal{P}$ (a graded poset) equals the length of any unrefinable chain from $\hat{0}$ to x.

We say that b *covers* a denoted $a \prec b$ if we have only the chain $\{a, b\}$ between a and b.

Every finite poset $\mathcal{P}$ has an associated simplicial complex $\Delta(\mathcal{P})$ called its *order complex*. For those not accustomed to the language of simplicial complexes, we now give a brief account.

Definition 6.2.5. Fix $n > 0$ an integer. A *k-simplex* Δ denoted $[i_1, \ldots, i_{k+1}]$ means a size $k+1$ subset $\{i_1, \ldots, i_{k+1}\} \subset \{1, \ldots, n\}$.

Definition 6.2.6. Call an r-simplex an *r-face* Δ' of a k-simplex Δ if $\Delta' \subset \Delta$.

Definition 6.2.7. A *simplicial complex* K means a finite family of simplices $\Delta \subset \{1, \ldots n\}$ such that if $\Delta_1 \in K$ and $\Delta_2 \subset \Delta_1$ then $\Delta_2 \in K$.

We call a simplicial complex K *k-dimensional* if the maximum dimension over every simplex in K equals k. In the case where every simplex occurs as the face of a k-simplex we call K *pure*.

Definition 6.2.8. To every finite poset $\mathcal{P}$ we associate the simplicial complex $\Delta(\mathcal{P})$ with simplices given by chains in $\mathcal{P}$.

To every simplicial complex we can associate a topological space called its *carrier* or *geometric realization*. To define this space we first define the associated geometric notions of simplices.

Definition 6.2.9. We call $k+1$ vectors $\{x_1, \ldots, x_{k+1}\}$ *affinely independent* exactly when we have the k vectors $\{x_2 - x_1, \ldots, x_{k+1} - x_1\}$ linearly independent.

Definition 6.2.10. A *geometric k-simplex* $\Delta \subset \mathbb{R}^m$ denoted $[x_1, \ldots, x_{k+1}]$ means the convex hull of $k+1$ affinely independent points in $\mathbb{R}^m$. Thus

$$\Delta = \operatorname{Conv}\left(\{x_i\}_{i=1}^{k+1}\right).$$

Call k the dimension of the simplex Δ.

Definition 6.2.11. An *r-face* of a geometric k-simplex $\Delta = \mathrm{Conv}\left(\{x_i\}_{i=1}^{k+1}\right)$ means an r-simplex occurring as the convex hull of $r+1$ points in $\{x_i\}_{i=1}^{k+1}$.

Fix a k-dimensional simplicial complex K. Fix a geometric $(n-1)$-simplex $\Delta = [x_1, \ldots, x_n] \subset \mathbb{R}^m$. Label the vertices of Δ as $\{1, \ldots, n\}$. For each simplex $\Delta' = [i_1, \ldots, i_k] \in K$ identify Δ' with the a copy $\overline{\Delta'}$ of the geometric simplex $[x_{i_1}, \ldots, x_{i_k}] \subset \Delta$. We now define an equivalence relation $\sim$ on the disjoint union of all such geometric simplices arising from K. Given two simplices $\overline{\Delta_1}$ and $\overline{\Delta_2}$ we identify their common face $\overline{\Delta_1 \cap \Delta_2}$. A geometric realization of K, or carrier of K means a topological space,

$$\overline{K} = \left(\coprod_{\Delta_i \in K} \overline{\Delta_i}\right) \Big/ \sim$$

Here $\overline{K}$ is a geometric simplicial complex.

Definition 6.2.12. A *geometric simplicial complex* $\overline{K}$ is a set of geometric simplices such that:

1. Every face of a simplex in $\overline{K}$ is again a simplex in $\overline{K}$;

2. The intersection of two simplices in $\overline{K}$ equals a common face of the two.

Note: The way the equivalence relation was defined above by gluing common faces together guarantees the second condition for the geometric realization of a simplicial complex.

Proposition 6.2.13. *A bounded poset $\mathcal{P}$ has a contractible carrier $\overline{\Delta(\mathcal{P})}$.*

Proof. $\mathcal{P}$ bounded makes its geometric realization a cone over $\hat{0}$ (or $\hat{1}$). So we have the contraction $(t, x) \mapsto (1-t)x + t\hat{0}$. □

Unboundedness of the poset makes determining the topology of the carrier tricky. Shellability offers one particularly nice condition for determining the topology of a simplicial complex.

Definition 6.2.14. Call a pure k-dimensional simplicial complex K *shellable* if we can order its maximal faces $\{\Delta_1, \dots, \Delta_N\}$ so that for each $l = 1, \dots, N-1$

$$\left(\bigcup_{i=1}^{l} \Delta_i\right) \cap \Delta_{l+1}$$

equals a pure $(k-1)$-dimensional simplicial complex.

Every pure shellable simplicial complex has a deformation to a bouquet of spheres of the same dimension. In the language of topology

Theorem 6.2.15. *(Bjorner, 1984) Given a shellable pure k-dimensional simplicial complex K*

$$\overline{K} \simeq \bigvee_i S^k.$$

We call a poset $\mathcal{P}$ *shellable* if and only if it has a shellable order complex $\Delta(\mathcal{P})$. Determining the shellability of a given complex turns out $\mathcal{NP}$-complete, meaning we can verify quickly that an ordering gives a shelling, but we have no efficient general algorithm to determine if a given complex has a shelling. This leads us to seek classes of labellings on chains in $\mathcal{P}$ that will guarantee the shelling of its order complex. One such labeling is an L-labeling.

Given a finite poset $\mathcal{P}$ define $C(\mathcal{P}) = \{(a, b) \in \mathcal{P} \times \mathcal{P} \,|\, a \prec b\}$.

Definition 6.2.16. An *edge labeling* of $\mathcal{P}$ means a function $\lambda : C(\mathcal{P}) \to \Lambda$ with Λ another poset.

Definition 6.2.17. Call an unrefinable chain $a_1 \prec \cdots \prec a_n$ is *rising* if $\lambda(a_1, a_2) < \cdots < \lambda(a_{n-1}, a_n)$.

Definition 6.2.18. An *L-labeling* of a graded poset $\mathcal{P}$ means an edge labeling $\lambda : C(\mathcal{P}) \to \Lambda$ that satisfying

1. Between two comparable elements $a \leq b$ there exists a unique rising unrefinable chain $a = a_1 \prec \cdots \prec a_n = b$;

2. Referring to the unrefinable chain above, if $a \prec c$ and $c \neq a_2$ then $\lambda(a, a_2) < \lambda(a, c)$.

Theorem. *(Bjorner,1980) An L-labeling makes graded poset $\mathcal{P}$ shellable.*

We refer the reader to [7] for the proof of this theorem.

Proposition 6.2.19. *$\hat{\mathcal{P}}$ is shellable if and only if $\mathcal{P}$ is shellable.*

Proof. Every maximal face in $\Delta(\mathcal{P})$ contains $\hat{0}$ and $\hat{1}$ as vertices. Hence, removing $\{\hat{0}, \hat{1}\}$ from $\mathcal{P}$ always reduces the dimension of the intersection of the ith maximal face in any shelling order with the union of the previous maximal faces by exactly 2. □

The proposition above gives us a powerful tool for determining the shellability of unbounded posets. This tool will be very useful later on to determine the topology of a certain important subset of ultrametrics. We now apply these results on posets to the poset of partitions.

6.3 The Poset of Partitions Π_n

Definition 6.3.1. Given $P, P' \in \Pi_n$ partitions of the set $[n] = \{1, \ldots, n\}$, call P *finer* than P' if for any part $A \in P$ there exists a part $A' \in P'$ such that $A \subset A'$. Equivalently, call P' *coarser* than P.

Definition 6.3.2. For $P \in \Pi_n$, $|P|$ equals the number of parts (subsets of [n]) contained in P.

Lemma 6.3.3. *If $P, Q \in \Pi_n$ with P strictly finer than Q, then $|P| > |Q|$. More specifically, two parts of P merge together in Q.*

Proof. Strictly finer means P finer than Q but $P \neq Q$. So P has a part x properly contained in some part $y \in Q$. $z = y \setminus x$ intersects some part $w \neq x$ of P. We must have $w \subset y$ lest two parts of Q intersect non-trivially. Hence parts x and w in P merge in Q. So that $|P| > |Q|$. □

Proposition 6.3.4. *The set of partitions Π_n forms a graded poset $(\Pi_n, \leq)$ with $P \leq P'$ for P finer than P'. Furthermore, the discrete partition furnishes the initial partition $\hat{0}$ and the indiscrete partition the terminal partition $\hat{1}$.*

Proof. Π_n clearly has the structure of a poset. The discrete partition has no refinement, the indiscrete partition no coarsening. Moving up a chain from discrete to indiscrete, the number of cells decreases, so chains have length at most n. We can't have a maximal chain shorter than n, or we'd have more than two cells merge in one step and we could refine the chain. □

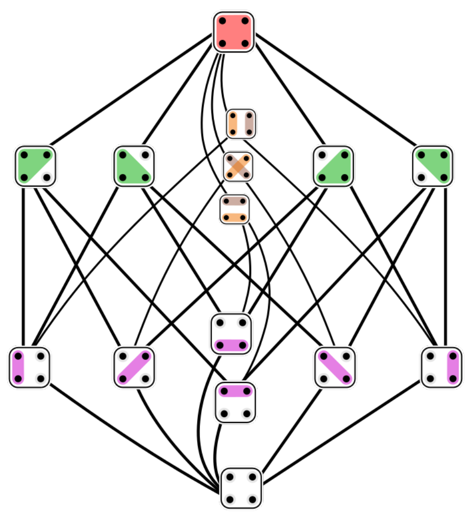


Figure 6.1: Lattice of Partitions of a 4-Element Set Tilman Piesk

Proposition 6.3.5 (Bjorner,1970)**.** Π_n *admits an L-labeling making both* Π_n *and* $\hat{\Pi}_n$ *shellable.*

Proof. Fix comparable partitions $P \leq Q$. If Q covers P, $P \prec Q$, we know that exactly two parts $x_{P,Q}$ and $y_{P,Q}$ of P merge together in Q. Define the map $\lambda : C(\Pi_n) \to [n]$ as

$$\lambda(P, Q) = \max \{\min \{x_{P,Q}\}, \min \{y_{P,Q}\}\}$$

More generally, all unrefinable chains between comparable partitions P and Q have the same length by the Jordan-Dedekind condition. We proceed by induction the length k of unrefinable chains between P and Q. For the basis we choose $k = 2$, we case $k = 1$ vacuous. Suppose $P \leq Q$ and assign to each part x of P the number $n_x = \min\{x\}$. Then define the set

$$A_{P,Q} = \{(x, y) \in P \times P \,|\, x, y \text{ distinct and merge in } Q\}$$

Let (x, y) be the minimum in $A_{P,Q}$ with respect to (n_x, n_y) in the dictionary order. Then let R denote the partition formed from P by merging x and y. By construction $\lambda(P, R) < \lambda(R, Q)$. If R' is any other cover of P, the two parts merged from P to form R' will not satisfy the minimality condition above and hence $\lambda(P, R) < \lambda(P, R')$. In addition, since x and y must eventually merge in Q we have that $\lambda(P, R') > \lambda(R', Q)$ so that $P \prec R \prec Q$ is the unique unrefinable rising chain. This proves the base case.

Now assume λ satisfies the conditions for an L-labeling if P and Q sit at length k from one another. Now if P and Q are length $k+1$ from each other let x be the part of P which is maximum with respect to n_x. Then let y be the part in Q which contains x. Define the partition P_k to be the partition formed from Q by unmerging x from y. By induction there exists a unique unrefinable rising chain $\mathcal{C} : P = P_1 \prec \cdots \prec P_k$ such that if $P \prec R$ and $R \neq P_2$. Then $\lambda(P, P_2) < \lambda(P, R)$. Obtain chain $\mathcal{C}'$ from $\mathcal{C}$ by appending Q at the end. Since we chose P_k so that it contained the maximum part of P with respect to n_x, we have that $\lambda(P_{k-1}, P_k) < \lambda(P_k, Q)$. This makes $\mathcal{C}'$ an unrefinable rising chain. In addition, any unrefinable rising chain connecting P to Q must include P_k as its penultimate member. For if not, x will be merged at some previous point in the chain, contradicting that the chain is rising. Thus, given an unrefinable rising chain from P to Q we obtain an unrefinable rising chain from P to P_k. By the induction hypothesis this chain must equal $\mathcal{C}$, making $\mathcal{C}'$ the unique unrefinable rising chain from P to Q. The chain $\mathcal{C}'$ inherits the second condition of an L-labeling from the subchain $\mathcal{C}$. This completes the induction. □

Corollary 6.3.6. *The carrier of order complex $\Delta(\hat{\Pi}_n)$ is homotopy equivalent to a wedge of $(n-3)$-spheres.*

The numbers of spheres in the decomposition above can be found by computing the Euler-characteristic of $\Delta(\hat{\Pi}_n)$. In the literature one usually does this by computing the values of a certain Mobius function associated to Π_n, which amounts to counting maximal chains with strictly decreasing Jordan-Holder sequences [See [7]]. Below we provided new proof that avoids this by defining an involution on the set of faces of $\Delta(\hat{\Pi}_n)$.

Definition 6.3.7. The *Euler characteristic* $\chi(K)$ of a k-dimensional simplicial complex K is given by the formula

$$\chi(K) = \sum_{i=0}^{k} (-1)^i k_i$$

where k_i denotes the number of i dimensional faces in K.

Theorem 6.3.8 (Feldman, Kehoe, 2019)**.** *There exists a computable bijection on the faces of* $\Delta(\hat{\Pi}_n)$*. As a result,*

$$\chi(\Delta(\hat{\Pi}_n)) = (-1)^{n-1}(n-1)! + 1$$

Proof. Let $X = \{1, \ldots, n\}$. We proceed by induction on n. We take $n = 3$ as base for the induction due to the non-vacuity of the chains. Here chains have length 1 and consist of a single partition which must separate 3 elements into 2 parts. The 3 possible partitions of this form make $\Delta(\hat{\Pi}_3)$ a set of three points, $\chi(\tilde{U}_3) = 3$, as predicted.

In general, assume now that the formula holds for $n-1$. We define an involution J on the set of faces of $\Delta(\hat{\Pi}_n)$ so that J matches unfixed faces to faces of a different parity in the Euler characteristic formula; this reduces calculating the Euler characteristic to counting and signing the fixed points of J.

The faces of $\Delta(\hat{\Pi}_n)$ correspond to non-empty chains that omit both extreme partitions (discrete and indiscrete), the set of which naturally bijects with the set of all chains that do contain both extreme partitions, but not just. An involution on the latter set immediately transfers to one on the set of faces of $\Delta(\hat{\Pi}_n)$, so we describe J there. Let $S = \{n\}$ and have J fix all chains of partitions that either

1. contain a partition of size $n-1$ that does not feature S, so n lies in the only part of size 2;

2. consist solely of the two extreme partitions and the size 2 partition containing S.

Now given any other chain, $\mathcal{C} = \{P_i\}_i$, define the set

$$A_{\mathcal{C}} = \{P \in \Pi_n \mid S \in P, \exists i\, P_i \leq P \leq P_{i+1}\}$$

and let

$$\pi = \min_{|P|} \{A_{\mathcal{C}}\}$$

where $|P|$ denotes the number of parts of P. In words, π equals the minimum size partition which contains the singleton S that may fit into the chain $\mathcal{C}$. Define J on these chains to be

$$J(\mathcal{C}) = \begin{cases} \mathcal{C} \cup \pi & , \quad \pi \notin \mathcal{C} \\ \mathcal{C} \setminus \pi & , \quad \pi \in \mathcal{C} \end{cases}$$

In any case $J^2(\mathcal{C}) = \mathcal{C}$, hence J defines an involution. In particular, J gives a bijection on the set of faces of $\Delta(\hat{\Pi}_n)$. We can see that J either fixes a face or matches faces differing by one dimension. J effectively toggles the presence of a partition π obtainable from a partition in $\mathcal{C}$ by splitting off singleton S from whatever non-singleton part that contains S. Note that we cannot toggle the presence of the partition π for chains of type 2 above. Indeed, deleting the size 2 partition would yield the chain consisting solely of the extreme partitions; a chain that corresponds to no face of $\Delta(\hat{\Pi}_n)$.

We now classify all fixed points $\mathcal{C}$ of type 1 according n's partner n_0 in the 2-element part belonging to $n-1$ size partition in $\mathcal{C}$. We have exactly $n-1$ possibilities for n_0, making $n-1$ classes of chains. Since each class makes the same contribution to the calculation of the Euler characteristic, we will fix n_0 and consider only type 1 chains in the n_0 class.

Transform all these chains by first deleting n from every part in every partition in every chain where it occurs, purging empty cells that might result, and finally removing (every copy of) every extreme partition. We then obtain the set of all extreme-partition-free chains of partitions of the $n-1$ element set $X \setminus S$, including the empty chain. We have the following observations

1. Removing the point n from every partition in $\mathcal{C}$ will leave two copies of the discrete partition, both of which get removed.

2. Removing n will not cause any partitions in any type 1 chain $\mathcal{C}$ to collapse. Indeed, only the discrete partition separate n from n_0. We can think of $\{n, n_0\}$ as a "fat point."

3. If we start with a chain associated to a face of $\Delta(\hat{\Pi}_n)$, append each end with extreme partitions, find we have a type 1 chain, delete n from every part of every partition in said chain, and then remove extreme partitions, we then obtain a chain one partition shorter than the one with which we started.

4. Shortening chains by 1 changes the sign of the contribution to the Euler characteristic that the corresponding faces make.

5. The unique 3-partition chain consisting of the two extreme partitions and the one size $n-1$ partition having all the elements of $X \setminus \{n, n_0\}$ as singletons transforms to that chain consisting solely of extreme partitions.

From these observations we can calculate the Euler characteristic of $\Delta(\hat{\Pi}_n)$. Observations 1 and 2, tell us that we can bijectively match type 1 class n_0 k-faces of $\Delta(\hat{\Pi}_n)$ to $(k-1)$-faces of $\Delta(\hat{\Pi}_{n-1})$ for $k \geq 1$. Observations 3 and 4 tell us that fixed points of type 1 in the class of n_0 make the contribution

$$\begin{aligned}(-1)\chi(\Delta(\hat{\Pi}_{n-1})) + 1 &= (-1)\left((-1)^{n-2}(n-2)! + 1\right) + 1 \\ &= (-1)^{n-1}(n-2)!\end{aligned}$$

to the Euler characteristic of $\Delta(\hat{\Pi}_n)$. Finally, there are exactly $n-1$ classes of type 1 chains, and there is exactly one type 2 chain; registering as a 0-face of $\Delta(\hat{\Pi}_n)$. Hence,

$$\begin{aligned}\chi(\Delta(\hat{\Pi}_n)) &= (n-1)\left((-1)^{n-1}(n-2)!\right) + 1 \\ &= (-1)^{n-1}(n-1)! + 1\end{aligned}$$

This completes the proof of the theorem. □

Once we know the Euler-Characteristic the number of spheres in a wedge of spheres $W = \bigvee_{i=1}^{N} S^k$ with $k > 1$ will be equal to $(-1)^k (\chi(W) - 1)$

Corollary 6.3.9. *The carrier of order complex $\Delta(\hat{\Pi}_n)$ is homotopy equivalent to a wedge of $(n-1)!$ $(n-3)$-spheres.*

We now turn our attention back to the space of ultrametrics.

6.4 The Complex of Ultrametrics

Let P_d^t denote the partition associated to the threshold metric d_t for $d \in U_n$ and $t \in I_d$. We have the following theorem describing the structure of $\{P_d^t\}_{t\in I_d}$.

Theorem 6.4.1 (Feldman, Kehoe, 2019)**.** *The collection $\mathcal{C}_d = \{P_d^t\}_{t\in I_d}$ forms a chain of length $1 \le k \le n$ in Π_n where k is the number of non-unital distances of d and $P_d^t \le P_d^{t'}$ for $t \le t'$.*

Proof. We first prove that $\mathcal{C}_d$ is in fact a chain in Π_n. Consider $0 \le t \le t' \le 1$ and $A \in P_d^t$ with $x_0 \in A$, then there exists $A' \in P_d^{t'}$ such that $x_0 \in A'$. Given $x \in A - \{x_0\}$ we have that $d(x, x_0) < t \le t'$ thus $d_{t'}(x, x_0) = 0$ so that $x \in A'$. Hence, $A \subset A'$ so that $P_d^t \le P_d^{t'}$. This proves that $\mathcal{C}_d$ is a chain in Π_n.

Now, since $d \in U_n$ by Proposition (6.1.2) we have that d admits
$1 \le k \le n$ non-unital distances $0 = \delta_1 < \delta_2 < \cdots < \delta_k < 1$. Let δ_{k+1} be unity, which

may or may not be a distance of d. We claim that P_d^t is constant on the intervals (δ_i, δ_{i+1}) for $1 \le i \le k+1$. Equivalently, suppose that $t, t' \in (\delta_i, \delta_{i+1})$ then we can show that P_d^t and $P_d^{t'}$ define the same equivalence relation; calling their associated relations $\sim_t$ and $\sim_{t'}$ respectively. If $x \sim_t y$ then $d(x,y) < t < \delta_{i+1}$, but $d(x,y) \in \{\delta_1, \dots, \delta_k, \delta_{k+1}\}$ so that $d(x,y) = \delta_j$ for $j \le i$. Hence, $d(x,y) \le \delta_i < t'$ showing that $x \sim_{t'} y$. By symmetry of the argument, we obtain that P_d^t and $P_d^{t'}$ define the same equivalence relation. Thus $P_d^t = P_d^{t'}$ and P_d^t (t not fixed) is constant on (δ_i, δ_{i+1}). Not only is P_d^t piecewise constant, but it obtains distinct values on the intervals (δ_i, δ_{i+1}). Indeed, if $d(x,y) = \delta_j$ then $x \sim_{\delta_j} y$ but $x \nsim_{\delta_i} y$ for $i < j$. From this we deduce that $|\mathcal{C}_d| = k$, the number of non-unital distances of d. □

To any given ultrametric d we can now associate a chain of partitions $\mathcal{C}_d$, but we can go further. As it turns out, d can be realized as the unique convex combination of partition metrics induced by $\mathcal{C}_d$.

Proposition 6.4.2 (Feldman,Kehoe, 2019)**.** *Given $d \in U_n$, d can be written uniquely as a convex combination of the partition metrics induced by $\mathcal{C}_d$. Moreover, the coefficients in the convex combination are all positive.*

Proof. Let $\delta_1 < \cdots < \delta_k$ denote the non-unital distances of d, and let $\delta_{k+1} = 1$. Choose $t_i \in (\delta_i, \delta_{i+1})$ for $1 \le i \le k$. Consider $x_j, y_j \in X$ such that $d(x_j, y_j) = \delta_j$ then we obtain the equations,

$$\begin{aligned}\delta_j &= \sum_{i=1}^{k} \lambda_i d_{t_i}(x_j, y_j) \\ &= \sum_{i=1}^{j-1} \lambda_i\end{aligned}$$

for $1 \leq j \leq k-1$. In the case that d takes the value of unity, we obtain the equation,

$$
\begin{aligned}
1 &= \sum_{i=1}^{k} \lambda_i d_{t_i}(x_j, y_j) \\
&= \sum_{i=1}^{k} \lambda_i
\end{aligned}
$$

which is exactly the convexity condition. If d does not take the value of unity then we only have the previous set of equations plus the condition that the coefficients $\{\lambda_i\}_{i=1}^k$ satisfy the convexity constraint. In any case solving for coefficients amounts to solving the augmented matrix,

$$
\left(\begin{array}{cccc|c}
1 & 0 & \cdots & 0 & \delta_2 \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
1 & \cdots & 1 & 0 & \delta_k \\
1 & \cdots & 1 & 1 & 1
\end{array}\right)
$$

which has the solution,

$$
\lambda_j = \textstyle\sum_{i=2}^{j+1} (-1)^{i+j+1} \delta_i \quad \text{for} \quad j = 1, \ldots, k
$$

Since $0 = \delta_1 < \delta_2 < \cdots < \delta_k \leq 1$ we have that $\lambda_j > 0$, and this completes the proof. $\square$

From here one might wonder, if convex combinations of partitions in a chain always give rise to ultrametrics. Our next proposition gives this converse.

Theorem 6.4.3 (Feldman, Kehoe, 2019)**.** *Let $\mathcal{C} = \{P_i\}_{i=1}^k$ be a chain of partitions of length k and denote the associated set of partitions metrics as $\{d_i\}_{i=1}^k$, then we have following,*

1. *Any convex combination partition metrics d_i gives an ultrametric. i.e. Conv($\mathcal{C}$) $\subset U_n$.*

2. *$\{d_i\}_{i=1}^k$ forms an affinely independent set of points in $\mathbb{R}^m$*

3. *If* $d = \sum_i \lambda_i d_i$ *with* $\sum_i \lambda_i = 1$ *and* $\lambda_i > 0$ *for all* i *then each* P_i *is a threshold partition of* d*. Specifically,* $\mathcal{C}_d = \mathcal{C}$ *where* $P_i = P_d^t$ *with*

$$\sum_{j=1}^{i-1} \lambda_j < t < \sum_{j=1}^{i} \lambda_j$$

for $1 \leq i \leq k$*.*

Proof. We begin by proving 1. Let $\{\lambda_i\}_{i=1}^k \subset \mathbb{R}^{\geq 0}$ such that $\sum_{i=1}^k \lambda_i = 1$, we need to check that $d = \sum \lambda_i d_i$ is an ultrametric. It is sufficient to check that d satisfies the strong triangle inequality. We'll show this in two steps. First consider $x, y \in X$. Since $\mathcal{C}$ forms a chain of partitions we have that if $d_i(x,y) = 1$, then $d_j(x,y) = 1$ for all $j < i$. Similarly, if $d_i(x,y) = 0$, then $d_j(x,y) = 0$ for all $j > i$. Now let J_{xy} denote the maximum index so that $d_i(x,y) = 1$ and let $J = 0$ in the case that no such index exists. Then we have,

$$d(x,y) = \sum_{i=1}^{J_{xy}} \lambda_i \, .$$

Let $z \in X$ and assume, with out loss of generality, that $J_{xy} \geq J_{yz}$ then

$$\begin{aligned}
d(x,z) &= \sum_{i=1}^{k} \lambda_i d_i(x,z) \\
&\leq \sum_{i=1}^{k} \lambda_i \max\{d_i(x,y), d_i(y,z)\} \\
&= \sum_{i=1}^{J_{xy}} \lambda_i \max\{d_i(x,y), d_i(y,z)\} \\
&= \sum_{i=1}^{J_{xy}} \lambda_i \\
&= d(x,y) \\
&= \max\{d(x,y), d(y,z)\}
\end{aligned}$$

Hence d satisfies the strong triangle inequality, so that d is an ultrametric.

We now prove that $\{d_i\}_{i=1}^k$ form an affinely independent set of points in $\mathbb{R}^m$. Let $v_i = d_1 - d_i$ for $2 \le i \le k$, we must show that $\{v_i\}$ form a linearly independent set of vectors in $\mathbb{R}^m$. Suppose not, then it is necessary that

$$\sum_{i=2}^{k} \lambda_i d_i = \left(\sum_{i=2}^{k} \lambda_i \right) d_1$$

where $\lambda_{i_0} \neq 0$ for some $2 \le i_0 \le k$. Let $\lambda = \sum_{i=2}^{k} \lambda_i$ then,

$$\sum_{i=2}^{k} \frac{\lambda_i}{\lambda} d_i = d_1.$$

Now, since the partitions in $\mathcal{C}$ are distinct we have that $P_1 < P_2$. Thus, there exist $x, y \in X$ separated in P_1, but together in the same part in P_2. Hence, $d_i(x, y) = 0$ for all $i \ge 2$ so that,

$$\begin{aligned} 1 &= \sum_{i=2}^{k} \frac{\lambda_i}{\lambda} d_i(x, y) \\ &= 0 \end{aligned}$$

a contradiction.

Finally we prove 3. Suppose $d = \sum_i \lambda_i d_i$ with $\sum_i \lambda_i = 1$ and $\lambda_i > 0$ for all i. Then let $x, y \in X$. By 1,

$$d(x, y) = \sum_{i=1}^{J_{xy}} \lambda_i \tag{6.4.4}$$

Suppose

$$\sum_{j=1}^{i-1} \lambda_j < t < \sum_{j=1}^{i} \lambda_j. \tag{6.4.5}$$

Then we'll show that $P_i = P_d^t$. Let $\sim_i$ and $\sim_t$ denote the equivalence via P_i and P_d^t respectively. Suppose that $x \sim_i y$. Then $d_j(x,y) = 0$ for all $j \geq i$ so that,

$$\begin{aligned} d(x,y) &= \sum_{j=1}^{i-1} \lambda_j d_j(x,y) \\ &\leq \sum_{j=1}^{i-1} \lambda_j \\ &< t \end{aligned}$$

Hence $x \sim_t y$. Now if $x \sim_t y$ then,

$$\begin{aligned} d(x,y) &= \sum_{j=1}^{J_{xy}} \lambda_j \\ &< t \end{aligned}$$

By (6.4.5), $J_{xy} \leq i-1$, so that $d_i(x,y) = 0$ and hence $x \sim_i y$. We now prove that in fact $\mathcal{C}_d = \mathcal{C}$; we can do this by calculating the number of distances of d. By (6.4.4), there exist at least k non-unital distances (including zero). If $P_k = \{X\}$ then λ_k will not contribute to d in distance, and hence d has exactly k non-unital distances. If $P_k \neq \{X\}$, then d takes on unity and will still have exactly k non-unital distances. By Theorem (6.4.1), $\mathcal{C}_d$ is a chain of length k. Hence, $\mathcal{C}_d = \mathcal{C}$. This completes the proof.

□

We can now precisely describe the structure of U_n.

Theorem 6.4.6 (Feldman, Kehoe, 2019)**.** *The set of ultrametrics U_n is the carrier of the pure $(n-1)$-dimensional simplicial complex Π_n.*

Proof. We can embed the poset Π_n into $\mathbb{R}^{\binom{n}{2}}$ by identifying partitions with their associated partition metrics. By Theorem (6.4.3) the convex hulls of the associated geometric simplices will be ultrametrics. □

We now turn our attention to the primary object of interest, the subcomplex of scaled ultrametrics.

6.5 Scaled Ultrametrics

Just as one can scale a metric and have it remain a metric, one can similarly affinely scale an ultrametric and have it remain an ultrametric. We define affine scaling by $(a,b) \in \mathbb{R}^2$ by first identifying d in $\mathbb{R}^m$ with the natural coordinates and then scaling each component, i.e.

$$(a,b) \cdot d := (a \cdot d_{12} + b, a \cdot d_{13} + b, \ldots, a \cdot d_{n-1,n} + b)$$

Notice that affine scaling does not affect self zero distances (condition 1) above.

Proposition 6.5.1. *Let $(a,b) \in \mathbb{R}^2$ and $d \in U_n$ then $\tilde{d} = (a,b) \cdot d \in U_n$ as long as $0 \le \tilde{d} \le 1$.*

Proof. Assume $0 \le \tilde{d} \le 1$ then the only condition to check is the strong triangle inequality, condition 3 above. But of course for x, y, z distinct

$$\begin{aligned} a \cdot d(x,z) + b &\le a \cdot \max\{d(x,y), d(y,z)\} + b \\ &= \max\{a \cdot d(x,y) + b, a \cdot d(y,z) + b\} \end{aligned}$$

The rest of the cases are routine checks. □

Corollary 6.5.2. *Every non-constant ultrametric d admits a unique affine scaling (a,b) so that $\tilde{d} = (a,b) \cdot d \in U_n$ with $\min(d) = 0$ and $\max(d) = 1$. (Here the min is taken over $d \in \mathbb{R}^m$, in other words, taking the minimum over the non-trivial distances of d.)*

Proof. Let $m = \min(d)$ and $M = \max(d)$ then solve for a, b in the system

$$am + b = 0$$

$$aM + b = 1$$

so that $a = \frac{1}{M-m}$ and $b = \frac{-m}{M-m}$. Since $M > m$ we have that a, b are well defined and the affine scaling will preserve minima and maxima. The result follows now as a direct corollary. □

Let $\tilde{U}_n := \{d \in U_n \mid \min(d) = 0, \max(d) = 1\}$ define the set of *scaled ultrametrics*. If we let R denote the set of constant ultrametrics then we can equivalently define $\tilde{U}_n$ as $\tilde{U}_n = (U_n - R)/\mathbb{R}^2$ with orbits under the action of affine scaling being identified with the unique zero-one representative given above.

Theorem 6.5.3 (Feldman, Kehoe, 2019). *$\tilde{U}_n$ is the carrier of the pure subcomplex $\Delta(\hat{\Pi}_n)$ of codimension 2 sitting in $\Delta(\Pi_n)$, given by the convex hulls of chains of partitions in Π_n that do not contain either the discrete or indiscrete partition.*

Proof. Let $d \in \tilde{U}_n$ then by Proposition (6.4.2) d induces a chain of threshold partitions $\mathcal{C}_d$ such that d is the unique convex combination,

$$d = \sum_{P \in \mathcal{C}_d} \lambda_P d_P$$

where $\sum_{P \in \mathcal{C}_d} \lambda_P = 1$ and each $\lambda_P > 0$. Since d is a scaled ultrametric there exists distinct $x_I, y_I \in X$ such that $d(x_I, y_I) = 0$. Hence,

$$\begin{aligned} d(x_I, y_I) &= \sum_{P \in \mathcal{C}_d} \lambda_P d_P(x_I, y_I) \\ &= 1 \end{aligned}$$

So that $d_P(x_I, y_I) = 1$ for all $P \in \mathcal{C}_d$, and therefore no P in $\mathcal{C}_d$ can be the indiscrete partition. Similarly, no P in $\mathcal{C}_d$ can be the discrete partition either. Thus $d \in \overline{\Delta(\hat{\Pi}_n)} \subset \mathbb{R}^{\binom{n}{2}}$. It is readily seen that inclusion of the discrete or indiscrete partitions in a convex combination for an ultrametric will automatically disallow non-trivial zero distances or unital distances respectively. Hence, $\tilde{U}_n$ is the carrier of $\Delta(\hat{\Pi}_n)$ □

Corollary 6.5.4 (Feldman, Kehoe, 2019)**.** *$\tilde{U}_n$ is homotopy equivalent to a wedge of $(n-1)!$ $(n-3)$-spheres.*

6.6 Iterated Cycle Structures

We now use the proof technique developed to calculate the Euler-characteristic in the last section to calculate a sort of Euler characteristic of the space of iterated cycle structures. Given a set X, call an element of $\mathrm{Sym}(X)$ a *cycle structure* on X, and *non-trivial* unless equal to the identity permutation. Given $\sigma \in \mathrm{Sym}(X)$, let $|\sigma|$ denote its set of cycles, and $||\sigma||$ the number of cycles.

Definition 6.6.1. An *iterated cycle structure* (ICS) on X of length $m \leq |X|$ means a finite sequence $(\sigma_1, \sigma_2, \cdots, \sigma_m)$ such that

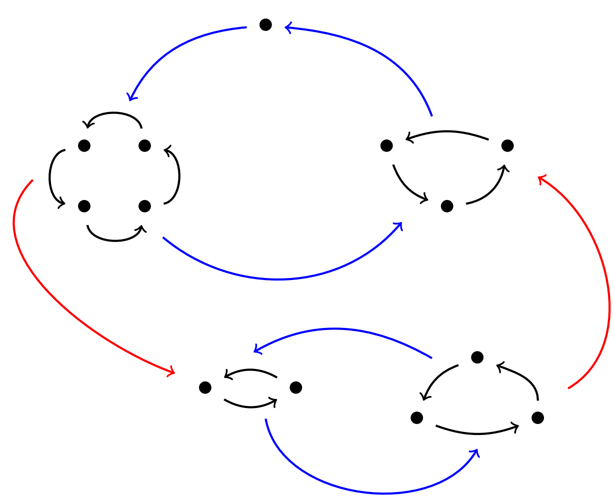
1) $\sigma_1 = \mathrm{id}_X$;

2) σ_{i+1} constitutes a non-trivial cycle structure on $|\sigma_i|$;

3) $||\sigma_m|| = 1$.

Write ICS(X) for the set of all iterated cycle structures on X.

Example 6.6.2. ICS of length 4 in ICS(**13**)



To each ICS $(\sigma_1, \sigma_2, \cdots, \sigma_m)$ we can associate *cycle-number sequence* $(||\sigma_1||, ||\sigma_2||, \cdots, ||\sigma_m||)$. By grouping members of ICS(X) according to their cycle-number sequence and summing over all possible cycle-number sequences we obtain the cardinality of ICS(X).

Let $\mathrm{St}(n, k)$ denote the Stirling numbers of the first kind. $\mathrm{St}(n, k)$ counts the number of permutations of n elements which are composed of k cycles. For for a cycle-number sequence $\{n_1, n_2, \ldots n_m\}$ there are exactly

$$\prod_{i=1}^{m-1} \mathrm{St}(n_i, n_{i+1})$$

choices of ICS. Rather than sum over all possible cycle-number sequences we can use matrix algebra to simplify our calculations.

Let $\mathbf{St}$ denote the n by n matrix over $\mathbb{Z}^+$ with entries given by,

$$\mathbf{St}(i, j) = \begin{cases} \mathrm{St}(i, j) & , \quad i > j \\ 0 & , \quad i \leq j \end{cases}$$

Proposition 6.6.3 (Feldman, Kehoe, 2019)**.** $\mathbf{St}^k(n, 1)$ *counts the number of ICS of length* $k+1$.

Proof. We first look at the case $k = 1$, here $\mathbf{St}(n, 1) = \mathrm{St}(n, 1)$; the number of ICS of length 2. For $k \geq 2$ we compute,

$$\begin{aligned} \mathbf{St}^k(n, 1) &= \sum_{j_1=1}^{n} \mathbf{St}^{k-1}(n, j_1)\mathbf{St}(j_1, 1) \\ &= \sum_{j_1, j_2=1}^{n} \mathbf{St}^{k-2}(n, j_2)\mathbf{St}(j_2, j_1)\mathbf{St}(j_1, 1) \\ &\vdots \\ &= \sum_{j_1, j_2, \ldots, j_{k-1}=1}^{n} \mathbf{St}(n, j_{k-1})\mathbf{St}(j_{k-1}, j_{k-2}) \cdots \mathbf{St}(j_2, j_1)\mathbf{St}(j_1, 1) \end{aligned}$$

By letting $j_0 = 1$ and $j_k = n$ we obtain

$$\mathbf{St}^k(n,1) = \sum_{j_1,\dots,j_{k-1}=1}^{n} \left(\prod_{i=0}^{k-1} \mathbf{St}(j_{i+1}, j_i) \right)$$

Using the fact that $\mathbf{St}(i,j) = 0$ for $i \leq j$ we have,

$$\begin{aligned}
\mathbf{St}^k(n,1) &= \sum_{1 \leq j_1 < \dots < j_{k-1} \leq n} \left(\prod_{i=0}^{k-1} \mathbf{St}(j_{i+1}, j_i) \right) \\
&= \sum_{1 \leq j_1 < \dots < j_{k-1} \leq n} \left(\prod_{i=0}^{k-1} \mathrm{St}(j_{i+1}, j_i) \right)
\end{aligned}$$

Each summand above is the number of ICS for a cycle-number sequence $(1, j_1, \dots, j_{k-1}, n)$. After summing these over all possible cycle-number sequences of length $k+1$ we obtain the desired result. □

Corollary 6.6.4 (Feldman, Kehoe, 2019)**.** $|\mathrm{ICS}(\mathbf{n})| = \sum_{k=1}^{n-1} \mathbf{St}^k(n,1)$

Write $\mathbf{n}$ for the set $\{1, 2, 3, \dots, n\}$. The numbers $|\mathrm{ICS}(\mathbf{n})|$ have received some attention; the sequence, goes

$$1, 1, 5, 47, 719, 16299, 513253, 21430513, 1145710573, \dots$$

Now attach weight $w_{n,m} = (-1)^{n-m}$ to each element $(\sigma_1, \sigma_2, \cdots, \sigma_m)$ of $\mathrm{ICS}(\mathbf{n})$.

Theorem 6.6.5 (Feldman, Kehoe, 2019)**.** *For all $n > 1$, the total of the weights on* $\mathrm{ICS}(\mathbf{n})$ *equals 1.*

Proof 1. We proceed with strong induction on n. For the base case we take $n = 1$. Here there is only the identity, the weight of which is 1. Now assume the result holds for $k < n$.

Fix $\sigma \in S_n$ and then consider the set,

$$S_\sigma = \{(\sigma_1, \sigma_2, \ldots, \sigma_m) \in \mathrm{ICS}(\mathbf{n}) \mid \sigma_2 = \sigma\}\,.$$

We can naturally identify S_σ with ICS($||\sigma||$) by treating σ_2 as an atom. Since $||\sigma|| < n$, by the induction hypothesis we have

$$\begin{aligned}
\sum_{(\sigma_1,\sigma_2,\ldots,\sigma_m)\in \mathrm{ICS}(S_\sigma)} w_{n,m} &= \sum_{(\sigma_1,\sigma_2,\ldots,\sigma_m)\in \mathrm{ICS}(S_\sigma)} (-1)^{n-||\sigma||} w_{||\sigma||,m-1} \\
&= (-1)^{n-||\sigma||+1} \sum_{(\sigma_1,\sigma_2,\ldots,\sigma_m)\in \mathrm{ICS}(S_\sigma)} w_{||\sigma||,m-1} \\
&= (-1)^{n-||\sigma||+1} \cdot 1 \\
&= (-1)^{n-||\sigma||+1}
\end{aligned}$$

Finally we must compute the sum of these weights over all $\sigma \in S_n$ sans the identity. To do this we'll simply calculate the whole sum and then adjust for the identity after.

It is a well-known fact in combinatorics that $\sum_{k=1}^{n}(-1)^k \mathrm{St}(n,k) = 0$. Hence,

$$\begin{aligned}
\sum_{\sigma\in S_n} (-1)^{n-||\sigma||+1} &= \sum_{k=1}^{n} \left(\sum_{||\sigma||=k} (-1)^{n-k} \right) \\
&= \sum_{k=1}^{n} (-1)^{n-k} \mathrm{St}(n,k) \\
&= (-1)^n \sum_{k=1}^{n} (-1)^k \mathrm{St}(n,k) \\
&= 0
\end{aligned}$$

To finish, adjust for the omission of the identity, (which would have counted -1), getting total weight 1, as desired. □

We now give an alternative proof of Theorem (6.6.5) by using a similar technique for the computation of the Euler-characteristic of the complex of scaled ultrametrics.

Proof 2. The standard order on $\mathbf{n}$ induces an order on every $|\sigma_i|$ in every ICS of the form $(\sigma_1, \sigma_2, \cdots, \sigma_m)$ on $\mathbf{n}$: to compare cycles in $|\sigma_i|$, using the induced order on $|\sigma_{i-1}|$, find and then compare their minimal elements.

For cycle structure σ with cycles $s_1, s_2, s_3, \ldots$ in increasing order define the prime of σ to be the cycle structure $\sigma' = ((s_1 s_2), s_3, \ldots)$ on $|\sigma|$. Let $\mathrm{ICS}^*(\mathbf{n})$ be the complement of the ICS $(\mathrm{id}_{\mathbf{n}}, (12), ((12)3), (((12)3)4), ...)$. For $(\sigma_1, \ldots \sigma_m) \in \mathrm{ICS}^*(\mathbf{n})$ let j denote the last index where $\sigma_{j+1} \neq \sigma_j'$. Then precompose σ_{j+1} with (12). One of two possibilities occur:

1. 1 and 2 are in the same cycle s_1 that break into two distinct cycles s_1 and s_2 in the new ordering.

2. 1 and 2 are in different cycles s_1 and s_2 that merge into one cycle s_1 in the new ordering

Define the involution $J : \mathrm{ICS}^*(\mathbf{n}) \to \mathrm{ICS}^*(\mathbf{n})$ by sending $(\sigma_1, \ldots, \sigma_m)$ to the ICS which has σ_{j+1} replaced with $\sigma_{j+1} \circ (12)$ and the remaining cycle structures the successive primes of $\sigma_{j+1} \circ (12)$. In case 1 above priming $\sigma_{j+1} \circ (12)$ will merge s_1 and s_2 back into a cycle s_1 in the new ordering. By the maximality condition on j, priming after this point will yield an ICS which agrees with $(\sigma_1, \ldots, \sigma_m)$. Thus the involution J effectively removes a cycle structure from $(\sigma_1, \ldots, \sigma_m)$ in case 1 or adds a cycle structure to $(\sigma_1, \ldots, \sigma_m)$ in case 2. In either case we pair off members of $\mathrm{ICS}^*(\mathbf{n})$ which are of the opposite parity according to their weight. Hence the total of the weights over $\mathrm{ICS}(\mathbf{n})$ will just be the weight of the ICS $(\mathrm{id}_{\mathbf{n}}, (12), ((12)3), (((12)3)4), ...)$ which is 1. □

Definition 6.6.6. A *pointed ICS* on X means an ICS $(\sigma_1, \sigma_2, \cdots, \sigma_m)$ with one marked cycle structure $\sigma_i, 1 < i \leq m$

Write $\mathrm{ICS}_+(X)$ for the set of pointed ICS's on X. As before we attach weight $(-1)^{n-m}$ to each element $(\sigma_1, \sigma_2, \cdots, \sigma_m)$ of $\mathrm{ICS}(\mathbf{n})$. Note that the choice of distinguished cycle does not affect the weight.

Write B_n for the n^{th} Bell number, the number of partitions of a set with cardinality n.

Theorem 6.6.7 (Feldman, Kehoe, 2019)**.** *For all $n > 1$, the total of the weights on $\mathrm{ICS}_+(\mathbf{n})$ equals $B_n - 1$. As a variation, define subset $\mathrm{ICS}_\oplus(X) \subset \mathrm{ICS}_+(X)$ by requiring that an ICS's final cycle* not *get marked. Then for all $n > 1$, the total of the weights on $\mathrm{ICS}_\oplus(\mathbf{n})$ equals $B_n - 2$. Indeed consideration of only those elements of $\mathrm{ICS}_+(\mathbf{n})$ that do have their final cycle marked brings us back to the previous theorem, so subtracting them all away reduces the weight by 1.*

Before we give a proof of this theorem we prove a useful lemma relating Bell numbers and Stirling numbers of the first kind

Lemma 6.6.8 (Feldman, Kehoe, 2019)**.**

$$B_n - 1 = \sum_{k=1}^{n-1} B_k \cdot \mathrm{St}(n,k) \cdot (-1)^{n-k+1}$$

or equivalently

$$1 = \sum_{k=1}^{n} B_k \cdot \mathrm{St}(n,k) \cdot (-1)^{n-k} \ .$$

Proof. $B_k \cdot \mathrm{St}(n,k)$ counts all partitions on all sets of cycles in cycle structures with k cycles. Excluding the choice of discrete partition of the identity cycle structure, let π be such a partition of a cycle structure σ. The partition π naturally induces a partition $\pi_{\mathbf{n}}$ on the set $\mathbf{n}$ by forgetting the cycle structure (just look at the points the cycle structure covers in every part). Since π is not a partition of the identity cycle structure there must be a part in π whose corresponding part in $\pi_{\mathbf{n}}$ contains at least 2 points. We'll call such a part a "fat" part.

Now from all fat parts in π identify in π the fat part P which contains the cycle with minimum element a of $\mathbf{n}$. Let b be the next largest element of $\mathbf{n}$ that is an element of any cycle in P. Note: a and b only depend on $\pi_{\mathbf{n}}$ and not on the cycle structure. Let σ' denote the cycle structure $\sigma \circ (ab)$. Just as in the Proof 2 of Theorem (6.6.5), $\sigma \circ (ab)$ will change the number of cycles by a factor of 1, thereby changing the parity of the weight $(-1)^{n-k}$. Define $J(\pi)$ to be the partition of the cycle structure σ' such that $J(\pi)_{\mathbf{n}} = \pi_{\mathbf{n}}$. Then J is a parity switching involution and thus matches partitions of different parities in the right hand side of the claimed identity above. Since we have excluded the discrete partition of the identity cycle structure we obtain,

$$0 = \left(\sum_{k=1}^{n} B_k \cdot \mathrm{St}(n,k) \cdot (-1)^{n-k} \right) - 1$$

The lemma is proved. □

Proof of Theorem (6.6.7). Given a cycle structure σ, ICSs in $\mathrm{ICS}_+(n)$ with $\sigma_2 = \sigma$, will either have σ_2 marked, or not.

Collecting together ICSs with σ_2 marked, seeing as we have no further cycle structure marked, we may appeal to the previous theorem by treating the cycles of σ_2 as atoms. These ICSs thus contribute a total weight of $(-1)^{n-||\sigma||+1}$.

We now proceed with strong induction on n. Taking as base case $n = 2$, we have one pointed ICS (the marked transposition ICS) and it has weight $B_2 - 1 = 1$. Collecting together ICSs with σ_2 unmarked, and using the induction hypothesis, we may assume we have total weight, just of those elements in $\mathrm{ICS}_+(n)$ with unmarked $\sigma_2 = \sigma$ equal to $(-1)^{n-||\sigma||+1} \cdot (B_{||\sigma||} - 1)$.

Putting the two cases together gives the total weight from ICSs with $\sigma_2 = \sigma$ as $(-1)^{n-||\sigma||+1} \cdot (B_{||\sigma||})$.

By Lemma (6.6.8) we have that the sum of weights over $\mathrm{ICS}_+(\mathbf{n})$ is

$$\sum_{k=1}^{n-1} (-1)^{n-k+1} \cdot \mathrm{St}(n,k) \cdot (B_k) = B_n - 1$$

The theorem is proved. □

Proof 2. We give a bijective proof of the theorem in its variant form. Call all partitions of a set other than the discrete and indiscrete *non-trivial.* Assuming that finite set X carries a total order, we build an involution I_X on $\text{ICS}_\oplus(X)$ with manifestly weight-canceling 2-cycles and positively-weighted fixed points that bijectively code non-trivial partitions of X. Specifically, we obtain $I_\mathbf{n}$.

The structure of I_X must certainly reflect the manner of coding partitions, but rather than exhibiting a specific I_X, we describe a general recipe for constructing I_X relative to a broad class of coding strategies. We can work with any method of encoding non-trivial partitions of a given ordered set X as elements of elements of $\text{ICS}_\oplus(X)$ provided it meets the following stipulations:

(i) code elements (for non-trivial partitions of X) have σ_2 marked;

(ii) code elements have length $|X|$ (and hence positive weight); and

(iii) the element ω generated by priming all the way occurs as a code.

Aside from these particulars, the description of I_X will stand indifferent to coding particulars.

By (ii), every cycle structure σ_i, $i \geq 2$ in every code element possesses, aside from 1-cycles, one single 2-cycle. Note that since $B_2 - 2 = 0$ and $B_3 - 2 = 3 = 6!/2$, the code space indeed has sufficient capacity.

In the spirit of the previous proof, the definition of $I_\mathbf{n}$ begins by separating two cases according to whether or not an an ICS has σ_2 marked.

Included within the set of elements with marked σ_2, by the stipulation above, sit all the code elements. $I_\mathbf{n}$ must leave code elements fixed. Nevertheless, we now utilize the involution specified in the second proof of the previous theorem. Recall, this involution fixed just a single element, namely ω, and we've stipulated that ω codes a partition. Thus $I_\mathbf{n}$ will ultimately also leave it fixed. However where the old involution paired other code elements with particular negatively-weighted, σ_2-marked elements of $\text{ICS}_\oplus$, we shall call these negatively-weighted elements level 2 *shadow codes*, and (for now) leave the behavior of $I_\mathbf{n}$ undefined both on code elements and their shadow codes. Note that the old fixed point has no shadow, so the number of shadow codes equals $B_n - 3$.

For ICSs with σ_2 unmarked, we iterate the procedure in the previous paragraph. Specifically for each cycle structure σ on $\mathbf{n}$, we apply the involution from the previous theorem to those ICSs with $\sigma_2 = \sigma$ and σ_3 marked, generating level 3 shadow codes; then with σ_2 and σ_3 specified and σ_4 specified and marked, for level 4 shadow codes. Carried to completion, this phase of the involution definition leaves as still unspecified the involution's behavior on two classes of elements of $\text{ICS}_{\oplus}$:

(i) permutation codes at every level; and (ii) shadow codes of every level. To deal with these, we start with just level 2 and level 3.

Now recall how in the previous proof we used transpositions to construct an involution verifying

$$1 = \sum_{k=1}^{n} B_k \cdot \text{St}(n,k) \cdot (-1)^{n-k} \ .$$

This involution preserves the number of cells in a partition even as it changes the cycles the cells contain. In particular, the involution still remains well-defined upon dropping all indiscrete (1-celled) partitions from the story. So already we have an involution that proves

$$1 = \sum_{k=1}^{n} (B_k - 1) \cdot \text{St}(n,k) \cdot (-1)^{n-k}$$

or

$$B_n - 2 = \sum_{k=1}^{n-1} (B_k - 1) \cdot \text{St}(n,k) \cdot (-1)^{n-k-1} \ .$$

To apply this, we now introduce into the mix partitions on the cycles of the identity permutation, encoded as usual, meaning as ICSs with σ_2 (not σ_3) marked followed by priming. Never mind for the moment that we have already defined the involution on these! The transposition involution matches

(i) certain partitions of $n-1$ cycles with

(ii) these partitions of n cycles (1-cycles!).

Now we fix everything by rematching the partitions in (i) with the previous partners of the partitions in (ii), and thus we liberate the partitions in (ii) to serve as the desired fixed points. Of course $B_n - 2$ partitions together with ω give the desired result. $\square$

LIST OF REFERENCES

[1] D. Avis, *On the extreme rays of the metric cone*, Adv. Math **no. 1** (1992), 47–105.

[2] M. Beck and S. Robins, *Computing the Continuous Discretely*: *Integer-Point Enumeration in Polyhedra*, Springer, 2009.

[3] D. Berrend and T. Tassa, *Improved bounds on Bell numbers and on moments of sums of random variables*, Probability and Mathematical Statistics **no. 30** (2010), 185–205.

[4] J.A. Bondy and U.S.R. Murty, *Graph Theory*, Springer, 2008.

[5] L Brickman, *Mathematical Introduction to Linear Programming and Game Theory*, Springer-Verlag, 1989.

[6] A. Bjorner, *Some Combinatorial and Algebraic Properties of Coexeter Complexes and Tits Buildings*, Advances in Mathematics (1970), 173–212.

[7] ______, *Shellable and Cohen-Macaulay Partially Ordered Sets*, Transactions of the American Mathematical Society **260** (1980), 159–183.

[8] P. Buneman, *The recovery of trees from measures of dissimilarity*, Mathematics in the Archaelogical and Historical Sciences (1971), 387–395.

[9] M.M. Deza and M. Laurent, *Geometry of Cuts and Metrics*, First Edition, Springer-Verlag Berlin Heidelberg, 1997.

[10] A. Dress and H-J. Bendelt, *A canonical decomposition theory for metrics on a finite set*, Canad. J. Math. 32 **no. 1** (1992), 126–144.

[11] X. Goaoc, P. Patak, Z. Patakova, M. Tancer, and U. Wagner, *Shellability is NP-complete* (2018), available at `https://arxiv.org/abs/1711.08436`.

[12] Branko Grunbaum, *Convex Polytopes*, Second Edition, Springer, 2003.

[13] Allen Hatcher, *Algebraic Topology*, Cambridge Press, 2001.

[14] J. Koolen, V. Moulton, and U. Tounges, *A classification of* 6*-point prime metrics*, European J. Combin. 21 **no. 6** (2000), 815–829.

[15] J. Spencer, *Maximal consistent families of triplets*, J. Combinatorial Theory (1968), 1–8.

[16] H. Wilf, *generatingfunctionology*, Springer, 2008.