DEGREE PROJECT IN MATHEMATICS,
SECOND CYCLE, 30 CREDITS
STOCKHOLM, SWEDEN 2021

Voronoi Cells of Varieties with respect to Wasserstein Distances

ADRIAN BECEDAS

KTH ROYAL INSTITUTE OF TECHNOLOGY
SCHOOL OF ENGINEERING SCIENCES

Voronoi Cells of Varieties with respect to Wasserstein Distances

ADRIAN BECEDAS

Master in Mathematics
Date: May 31, 2021
Supervisor: Kathlén Kohn
Examiner: Kathlén Kohn
Royal Institute of Technology

Abstract

Voronoi diagrams are partitions of a metric space into Voronoi cells according to distance from points on some set w.r.t. some distance. In this thesis we examine Voronoi diagrams of manifolds and varieties w.r.t. the Wasserstein distance from probability theory. We give some upper and lower bounds on the dimension of Voronoi cells based on the geometry of the manifolds and Wasserstein distance balls. We provide an upper bound on the number of full-dimensional Voronoi cells of algebraic varieties and show examples of the bound being tight.

Sammanfattning

Voronoi diagram är partitioner av ett metriskt rum i Voronoi celler enligt avstånd från punkter på någon mängd med avseende på ett visst mått. I den här avhandlingen undersöker vi Voronoidiagram för mångfald och varieteter med avseende på Wasserstein avstånd från sannolikhetsteori. Vi ger några övre och nedre gränser för dimensionen på Voronoiceller baserat på geometrin hos mångfalden och avståndsklot med Wasserstein mått. Vi presenterar en övre gräns för antalet fulldimensionella Voronoiceller för algebraiska varieteter och visar exempel på när gränsen är strikt.

Acknowledgements
I am deeply grateful to my supervisor, Kathlén Kohn, for her indispensable guidance and direction. I would also like to express my heartfelt thanks for the invaluable discussions and support of both Kathlén Kohn and Lorenzo Venturello.

Contents

Chapter 1

Introduction and Outline

In the second chapter of this thesis we introduce the concept of the Wasserstein distance. We give some historical background and define the Wassertein distance balls. We also provide some background on Voronoi diagrams and define Voronoi cells.

The third chapter is dedicated to understanding Voronoi cells w.r.t Wasserstein distances. We introduce cell-cones which allow us to break down Voronoi cells into more manageable components. We prove a series of results which produce upper and lower bounds for dimensions of Voronoi cells depending on the orientation of faces of the Wasserstein distance ball.

We spend the fourth chapter more thoroughly exploring the Voronoi diagrams of the statistically significant Hardy-Weinberg curve in the simplex w.r.t different Wasserstein metrics.

The final chapter is where we find an upper bound for the number of full-dimensional Voronoi cells for smooth irreducible varieties with hypersurfaces as dual varieties. We use dual varieties and polar degrees for the upper bound w.r.t. a generic Wasserstein distance. We also reference adjecent research and properties of generic distance balls and provide an example when our bound is sharp.

Chapter 2

Prerequisites

2.1 Wasserstein distance

According to *Encyclopedia of Mathematics* [9] the Wasserstein distance first appears in writing as "Vasershtein distance" in 1970 in a discussion on probability measures. The name then referred to a paper from 1969 by Vasershtein on Markov processes. However the earliest appearence of the distance was in 1940 in a paper by Kantorovich on transportation problems.

The Wasserstein distance also appears as a solution to the following optimization problem:

$$\text{maximize} \sum_{i=1}^{n} (\mu_i - \omega_i) x_i$$

subject to $|x_i - x_j| < d_{i,j} \in \mathbb{R}$, for all $0 < i < j \leq n$, where μ, ω are probability distributions.

However for our purposes we will be satisfied with the following definition for distance balls of Wasserstein distances, on the subspace of $\mathbb{R}^{n+1}$ given by $\mathbf{1}_n = \{(t_1, \cdots, t_{n+1}) | \sum t_i = 1\}$.

Definition 2.1. *Let $D \in \mathbb{R}^{(n+1)\times(n+1)}$ be a matrix satisfying*

$$d_{i,j} = d_{j,i}, d_{i,i} = 0, d_{i,j} \leq d_{i,k} + d_{k,j}.$$

The Wasserstein associated to D has distance balls $B_{D,r}(c)$ *of radius $r \in \mathbb{R}$ centered at $c \in \mathbf{1}_n$ defined as the convex hull of the set*

$$\{c + r\frac{u_i - u_j}{d_{i,j}} | i \neq j\},$$

where u_i is the i'th basis vector in $\mathbb{R}^{n+1}$.

For simplicity we also introduce the following notation.

Notation 2.2. *If the position and radius of a Wasserstein distance ball $B_{D,r}(c)$ are not relevant we drop the D and r and write only B_D. We will use F to denote facets of distance balls and f for faces of any dimension. Since we will want to talk about the symmetric faces of distance balls we will introduce the shorthand $-f$ for the opposite face spanned by points $\{c - r\frac{u_i - u_j}{d_{i,j}}\}$ when the face f is spanned by points $\{c + r\frac{u_i - u_j}{d_{i,j}}\}$.*

Let us first look at some examples of Wasserstein distance balls and their associated matrices. The following are examples of balls in $\mathbf{1}_2$, identified with $\mathbb{R}^2$, compared with the 2-simplex, $\Delta_2 := \{(t_1, t_2, t_3) \in \mathbb{R}^3_{\geq 0} | t_1 + t_2 + t_3 = 1\}$.

Example 2.3. *The matrix*

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

together with radius $\frac{1}{3}$ and the barycenter of the simplex as center of the ball gives us the distance ball in Figure 2.1.

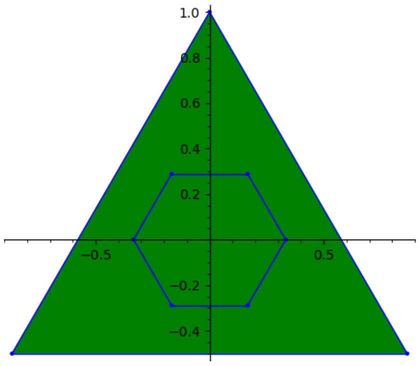

Figure 2.1: Distance ball from Example 2.3 in simplex.

Example 2.4. *With the same choices of radius and center as the last and*

$$\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}$$

as the associated matrix we obtain the ball in Figure 2.2.

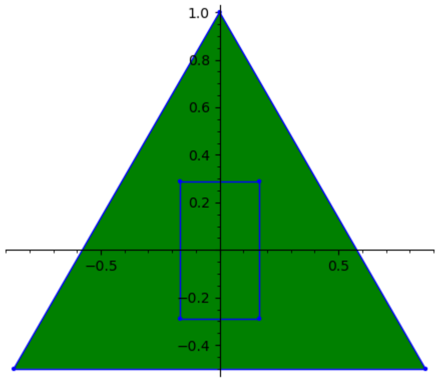

Figure 2.2: Distance ball from Example 2.4 in simplex

These two examples presented show all cases for the possible number of sides of the Wasserstein distance balls in $\mathbf{1}_2$ which we will show in the following proposition.

Proposition 2.5. *Wasserstein distance balls in* $\mathbf{1}_2$ *are polygons with six or four sides or, when the radius is 0, a single point.*

Proof. Since the Wasserstein distance balls $B_{D,r}(c)$ in $\mathbf{1}_2$ are the convex hulls of 6 points in a plane they can only be 6,5,4 or 3 sided polygons, a line segment or a point. Because the 6 points are endpoints of evenly spaced vectors around c, with length $\frac{r}{d_{i,j}}$ the line segment can be ruled out and the point only achieved when r=0. The symmetry of the vectors $c + \frac{u_i - u_j}{d_{i,j}}, c + \frac{u_j - u_i}{d_{j,i}}$ means we have an even number of edges/vertices when $r \neq 0$. □

2.2 Voronoi diagrams

Voronoi diagrams are partitions of a metric space X into Voronoi cells. Given some set of points S and a distance d, the Voronoi cells are the set of points closer to one specific point in S than any other point in S.

The traditional Voronoi diagrams are ones where the metric space is $\mathbb{R}^2$ with the euclidean distance and S is a finite set of points. In this case all points s in S have an associated Voronoi cell which is a convex polygon with s as an interior point.

In *Spacial Tessellations* [4] we are told that Descarte wrote about diagrams much like Voronoi diagrams in the early 1600s for his writings on astronomy and light. It is however first in the 1900's that Voronoï, who the Voronoi diagrams got their names from, wrote his first papers on the subject. The reason for the name is said to be that Voronoïs papers were the earliest giving a comprehensive view of the diagrams as well as generalise the concept to the higher dimensional cases. Since the 1970s research on Voronoi diagrams has expanded to include diagrams for many different sets S as well as different distance functions. Areas of application include crystallography, meteorology, ecology, and code study. Perhaps the most famous example is their use in a report on Cholera in 1854.

We will mainly be interested in Voronoi diagrams with Wasserstein distance and where the set S is some smooth manifold or variety. Voronoi cells of varieties were first defined recently in *Voronoi cells of varieties* [5] which focused on the euclidean distance. Voronoi cells of varieties with Wasserstein distance are investigated first in this thesis.

Definition 2.6. *With the notation described above, the Voronoi cell of a point $s \in S$ is the set*

$$V_{d,S}(s) = \{x \in X | d(x, s) < d(x, s') \forall s' \in S, s \neq s'\}.$$

Chapter 3

Voronoi cells with respect to Wasserstein distances

In this section we discuss properties and conditions for Voronoi cells w.r.t. Wasserstein distances. First we define the concept of a cell-cone which is used to split Voronoi cells into components related to the faces of a given distance ball. In this section D is assumed to be a Wasserstein distance on $\mathbf{1}_n$ and f is assumed to be a face of the distance ball B_D. By $span(f)$ we mean the smallest affine space containing f.

Definition 3.1. *For some point $x \in \mathbf{1}_n$ and some face f of $B_{D,r}(x)$, the* cell-cone *of f at x is the set of points*

$$C_f(x) = \{x + \epsilon(s - x) | s \in -f, \epsilon > 0\}.$$

We also introduce the related notation

$$C_{f,\gamma}(x) = \{x + \epsilon(s - x) | s \in -f, \gamma > \epsilon > 0\}.$$

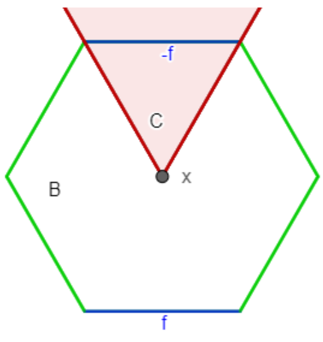

Figure 3.1: Example illustration of cell-cone from Definition 3.1.

The cell-cone is the set of all points such that the Wasserstein ball of some size around these points intersects x at face f, see Figure 3.1. We note that $C_{f,\epsilon}(x) \subset B_{D,\epsilon}(x)$. It can be noted that $C_f(x)$ does not depend on the radius of the ball $B_{D,r}(x)$ while $C_{f,\gamma}(x)$ does. Having the ball $B_{D,r}(x)$ with which $C_f(x)$ is defined be centered at the same point x as the cell-cone is not strictly necessary as the vector $s-x$ could be replaced with a vector from the center of any ball to a point on that balls face. The Voronoi cell for x can be partitioned into parts in the distinct relative interiors of cell-cones of the distinct faces of B_D. The following fact shows the usefulness of knowing in which cell-cone a point of the Voronoi cell lies.

Proposition 3.2. *Let X be a subset of the space $\mathbf{1}_n$ and let $x \in X$. Assume that $y \neq x$ is a point in the Voronoi cell $V_{D,X}(x)$ and let $\epsilon > 0$ be such that $B_{D,\epsilon}(y) \cap X = \{x\}$. If x is in the relative interior of the face f of $B_{D,\epsilon}(y)$, then the dimension of the Voronoi cell $V_{D,X}(x)$ is at least dim$(f) + 1$.*

Proof. Note that the points in the intersection $C_{-f,\epsilon}(y) \cap C_f(x)$ are in the Voronoi cell. This holds since for a point z in the intersection, there exists a $\delta > 0$ such that $x \in B_{D,\delta}(z) \subset B_{D,\epsilon}(y)$. To see this we let $\delta = D(z,x)$ and $p \in B_{D,\delta}(z)$. We then have $D(y,p) \leq D(y,z) + D(z,p) \leq (\epsilon - \delta) + \delta = \epsilon$ since $D(y,z) + D(x,z) = D(y,x) = \epsilon$.

Since the relative interior of the line segment from x to y is in both $C_{-f,\epsilon}(y)$ and $C_f(x)$, their intersection is a nonempty open subset in the dim$(f)+1$ dimensional subspace spanned by f and y. Hence the dimension of the Voronoi cell $V_{D,X}(x)$ is at least dim$(f) + 1$, the same dimension as the cell-cone. □

Proposition 3.2 shows that finding points in a Voronoi cell belonging to a certain cell-cone gives information about the dimension of the Voronoi cell. We now want a statement which gives a lower bound for the dimension of the Voronoi cell $V_{D,A}(x)$ depending on faces of the distance balls. We do this by splitting an open neighbourhood around x with a hyperplane which leaves X entirely on one side.

In order to do this, we introduce, for a face f of distance ball $B_{D,\epsilon}(c)$, the vector v_f. Let $v_f \in span(f \cup \{c\})$ be the vector that is orthonormal to f and points from c to f.

Under certain conditions we can think of the hyperplane $H_{f,x}$ passing through a point x defined as the orthogonal complement of v_f to separate a distance ball and the manifold x belongs to; see Figure 3.2.

Proposition 3.3. *Let X be a subset of the space $\mathbf{1}_n$, $x \in X$ and f be a face of some distance ball B_D. If there exists a radius $r > 0$ such that the Euclidean* *distance ball* $\beta_{r,x} \in \mathbf{1}_n$ *centered at x is bisected by* $H_{f,x}$ *and the closure of one of the resulting half-balls intersects* X *only at* x*, then the Voronoi cell* $V_{D,X}(x)$ *has dimension at least* $dim(f) + 1$*.*

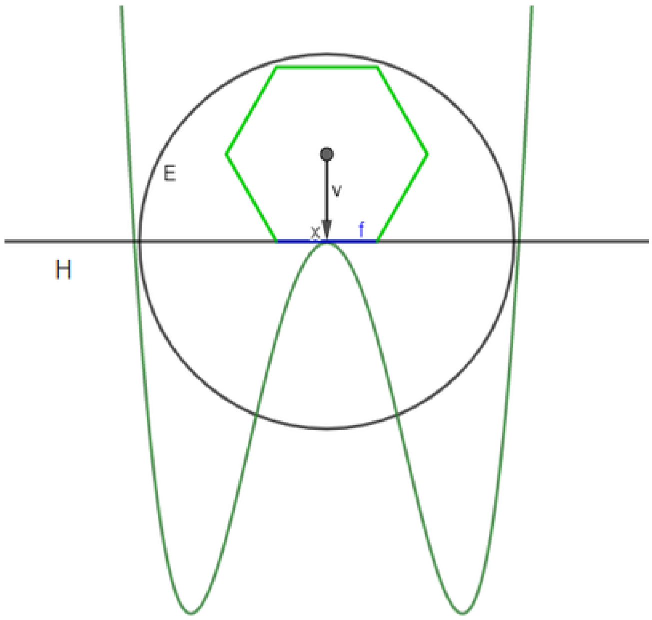

Figure 3.2: Example illustration of setup in proof of Proposition 3.3 in the plane.

Proof. Consider a Wasserstein distance ball $B_{D,\epsilon}(c)$ such that x is in the relative interior of f. We could choose the position of c and size of ϵ small enough that $B_{D,\epsilon}(c) \subset \beta_{r,x}$. By construction, $H_{f,x}$ splits $\beta_{r,x}$ into two open half-balls β' and β'' such that $B_{D,\epsilon}(c) \cap \beta' = \emptyset$. If β' is the half-balls from the statement, then $c \in V_{D,X}(x)$ and $c \in C_f(x)$. Otherwise a symmetric construction of $B_{D,\epsilon}(c)$ such that x is in the relative interior of $-f$ will give us such a c. By Proposition 3.2, $V_{D,X}(x)$ has dimension at least $dim(f) + 1$. □

Naturally we now want to find an upper bound for dimensions of Voronoi cells. To determine when we do not have large or full-dimensional Voronoi cells we would need knowledge of when no point in a Voronoi cell lies in a particular cell-cone. We now restrict our investigation of Voronoi diagrams to sets which are smooth manifolds in order to present a lemma to that effect. The next lemma gives a condition which when satisfied tells us that the Voronoi cell $V_{D,A}(x)$ does not contain any point in the relative interior, $C_f^\circ(x)$, of the cell-cone $C_f(x)$. To not make the statement of the lemma exceedingly lengthy we explain the condition ahead of time. Assume we have a smooth manifold A and a point $x \in A$. Furthermore, let $y \in C_f^\circ(x)$ and $B_{D,\epsilon}(y)$ be the distance ball such that x is in the relative interior of the face f of $B_{D,\epsilon}(y)$. Our condition (C) for a plane P is that $x, y \in P$ and that there exists two non-parallel line segments $l_f, l_t \subset P$ such that x is in the relative interior of each line segment, l_f is in the face f of $B_{D,\epsilon}(y)$ and l_t is in the tangent space of A at x. This property might seem arbitrary but having the lemma formulated in this manner will make it easier to prove other statements about Voronoi cells later.

Lemma 3.4. *Let A be a smooth manifold, $x \in A$ a point on the manifold and f a face of $B_D(x)$. If for all $y \in C_f^\circ(x)$ there exists a plane P with property (C), then the Voronoi cell $V_{D,A}(x)$ does not contain any point in $C_f^\circ(x)$.*

Proof. A point y is in the Voronoi cell $V_{D,A}(x)$ if and only if for some non-negative real ϵ the intersection $B_{D,\epsilon}(y) \cap A = \{x\}$.

We assume for contradiction that there exists a point $y \in V_{D,A}(x) \cap C_f^\circ(x)$. Let P be such a plane satisfying (C) with line segments l_f, l_t. Since l_f and y span P, $l_t \subset P$ and x is in the relative interior of f, we can find a Euclidean ball $\beta \subset B_{D,\epsilon}(y)$ around some point on l_t. Let L be the set of line segments from x to a point in β. Since $B_{D,\epsilon}(y)$ is convex, all points on line segments in L are in $B_{D,\epsilon}(y)$. For some small angle $\theta > 0$, all lines through x with angle at most θ from l_t in any direction contain some line segment in L.

Since l_t is in the tangent space of A at x, there is some $\delta > 0$ such that for all $0 < r < \delta$ there is a point a on A at distance r from x with angle less than θ from l_t. Hence for δ small enough, a is contained in a line segment in L and so a is also in $B_{D,\epsilon}(y)$. This contradicts $B_{D,\epsilon}(y) \cap A = \{x\}$, so the Voronoi cell $V_{D,A}(x)$ does not contain any point in $C_f^\circ(x)$. ☐

If Lemma 3.4 is applied to all faces of high dimensions, the lemma can be used to give an upper bound of the dimension of the Voronoi cell, this time in the language of transversal intersections.

Theorem 3.5. *Let A be a smooth manifold and $x \in A$ be a point on the manifold. Assume that f is a face of some Wasserstein ball $B_D(x)$ intersecting A at x non-transversally. If no other face of larger dimension has this property then the Voronoi cell $V_{D,A}(x)$ is of dimension at most $dim(f) + 1$.*

Proof. Let f' be any face of some Wasserstein ball $B_D(c)$ intersecting A at x transversally. Since A has lower dimension than the ambient space, $\mathbf{1}_n$, f' cannot be a zero-dimensional face. Since f' intersects A at x transversally, any vector in the ambient space, $\mathbf{1}_n$, is the sum of a vector in the subspace spanned by f' and a vector in the tangent space of A at x. In particular, since the center of the ball c is not contained in f', the vector v_c from x to c is the sum of a vector v_f in f' and a vector v_t in the tangent space of A at x that is not in f'. The plane spanned by these three vectors satisfies property (C). We can move the center c to an arbitrary element of $C^\circ_{f'}(x)$ such that the new face f' still intersects A at x transversally. By Lemma 3.4 the cell-cone component $V_{D,A}(x) \cap C^\circ_{f'}(x)$ associated with f' is empty.

Recall that we can partition the Voronoi cell into the cell-cone components. All faces of dimension larger than $dim(f)$ intersect A at x transversally, so their corresponding cell-cone components are empty. We are left with a Voronoi cell contained in the union of finitely many cell-cones of faces of dimension dim(f) or lesser. We conclude that $V_D(x)$ is of dimension at most dim(f)+1. □

For a smooth manifold of codimension 1, the tangent space at any point is of codimension 1 and non-transversal intersection is reduced to tangency. This gives us the following corollary.

Corollary 3.6. *Let x be some point on a smooth manifold A of dimension one less than the ambient space. If the face f is the largest dimensional face of some Wasserstein ball such that $x \in f$ and f is contained in the tangent space of A at x then the Voronoi cell $V_{D,A}(x)$ is of dimension at most dim(f)+1.*

By similar reasoning we get the corollary

Corollary 3.7. *Let x be a point on a smooth manifold A. The Voronoi cell $V_{D,A}(x)$ is not full-dimensional unless the tangent space of A at x is contained in the hyperplane spanned by a facet F of some Wasserstein ball B_D.*

In the cases expressed in the corollaries, the condition of non-transversality reduces to tangency. We can now collect some equivalent properties for full-dimensional Voronoi cells when we have tangency at a facet in a lemma.

Corollary 3.8. *Assume we have some point x on a smooth manifold A and a facet F of a Wasserstein ball B_D such that x is in the relative interior of F (or $-F$) and the tangent space of A at x is contained in the hyperplane spanned by F (or then one spanned by $-F$). In this situation the following are equivalent:*

(1) W.r.t. the Wasserstein distance D, x is the closest point of A to the center c of some Wasserstein ball $B_{D,\epsilon}(c)$ such that x is in the relative interior of F or $-F$.

(2) There is some small neighborhood around x where the hyperplane spanned by F bisects the neighborhood such that all points of $A \setminus \{x\}$ lie strictly in one of the bisected components.

(3) The component of the Voronoi cell $V_{D,A}(x)$ associated with the cell-cone of F or $-F$ is full-dimensional, hence $V_{D,A}(x)$ is full-dimensional.

Proof. The proof ofProposition 3.2 gives us the implication (1) $\Longrightarrow$ (3).
The proof of Proposition 3.3 gives us the implication (2) $\Longrightarrow$ (3).
By the properties of the cell-cones, we know that (3) $\Longrightarrow$ (1).
Choosing a Euclidean ball around x small enough so that F or $-F$ from (1) bisects the ball with one bisected side contained entirely in $B_{D,\epsilon}(c)$ we see that (1) implies (2). □

Remark 3.9. *For arbitrary faces, we can replace the hyperplane in (2) by the hyperplane $H_{f,x}$ from Proposition 3.3 and "full-dimensional" in (3) with "$(dim(f) + 1)-$dimensional". In that case, (1) and (3) are equivalent and (2) implies (1) and (3) but not vice versa.*

Chapter 4

Voronoi diagrams of Hardy-Weinberg curve in the simplex with Wasserstein distance

Let us examine the full-dimensional Voronoi cells for a specific curve in the plane with different Wasserstein distances. We will choose the Hardy-Weinberg curve which is the 2-dimensional member of a family of statistically significant curves in their respective simplices. This family is known as the Veronese curves, which are parametrized by

$$(\binom{n}{i}p^i(1-p)^{n-i})_{i=0}^n \in \Delta_n$$

where $p \in [0,1]$ and $\Delta_n = \{(t_0, \cdots, t_n) \in \mathbb{R}_{\geq 0}^{n+1} | t_0 + \cdots + t_n = 1\}$.

Calculating the general tangent vectors for points on the curves gives us

$$(\binom{n}{i}ip^{i-1}(1-p)^{n-i} - \binom{n}{i}(n-i)p^i(1-p)^{n-i-1})_{i=0}^n.$$

The 2- dimensional Hardy-Weinberg curve therefore has the form $(p^2, 2p(1-p), (1-p)^2)$ with tangents given by $(2p, 2-4p, 2p-2)$. We note that the tangent vectors vary linearly from $(0,2,-2)$ to $(2,-2,0)$ and that these two vectors are parallel to the sides of the simplex at the respective corners which are the curve at respective endpoints. Armed with Corollaries 3.8 and 3.6 we investigate three Wasserstein distances by their matrices.

Our first Wasserstein distance is given by the following matrix $M_1 =$

$$\begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 0 \end{pmatrix}$$

and has sides with direction vectors (1,0,-1),(1,1,-2) and (2,-1,-1). Examining the curve, we see that these vectors are tangent vectors of the curve at the following points:

p	$(p^2, 2p(1-p), (1-p)^2)$	$(2p, 2-4p, 2p-2)$
$\frac{1}{2}$	$(\frac{1}{4},\frac{1}{2},\frac{1}{4})$	(1,0,-1)
$\frac{1}{3}$	$(\frac{1}{9},\frac{4}{9},\frac{4}{9})$	$(\frac{2}{3},\frac{2}{3},-\frac{4}{3})$
$\frac{2}{3}$	$(\frac{4}{9},\frac{4}{9},\frac{1}{9})$	$(\frac{4}{3},-\frac{2}{3},-\frac{2}{3})$

Corollary 3.6 tells us that these are the only points which can have full-dimensional Voronoi cells and Corollary 3.8 confirms that these are points with full-dimensional Voronoi cells. Below is an illustration of the curve in the simplex with a centred distance ball corresponding to the matrix.

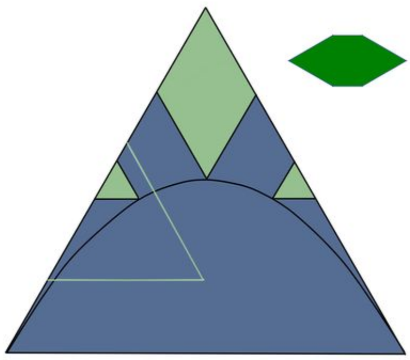

Figure 4.1: Voronoi cells (light green) of Hardy-Weinberg curve and a distance ball (top right) associated with M_1 in the simplex.

Our next example is of a matrix M_2, shown below, giving rise to tangency at two points.

$$\begin{pmatrix} 0 & 8 & 3 \\ 8 & 0 & 11 \\ 3 & 11 & 0 \end{pmatrix}$$

In this case this happens because one of the inequalities governing the matrices is strict and so some points in Definition 2.1 are co-linear. This reduces the number of sides of the distance ball as can be seen in Figure 4.2 . The two direction vectors are $(\frac{5}{24}, \frac{1}{8}, -\frac{1}{3})$ and $(\frac{11}{24}, -\frac{1}{8}, -\frac{1}{3})$. These vectors are tangent vectors of the curve at the following points:

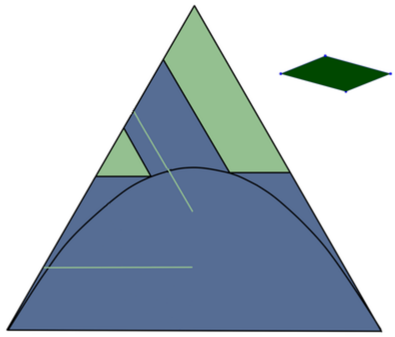

Figure 4.2: Voronoi cells (light green) of Hardy-Weinberg curve and a distance ball (top right) associated with M_2 in the simplex.

p	$(p^2, 2p(1-p), (1-p)^2)$	$(2p, 2-4p, 2p-2)$
$\frac{5}{13}$	$(\frac{25}{169}, \frac{80}{169}, \frac{64}{169})$	$(\frac{10}{13}, \frac{6}{13}, -\frac{16}{13})$
$\frac{11}{19}$	$(\frac{121}{361}, \frac{176}{361}, \frac{64}{361})$	$(\frac{22}{19}, -\frac{6}{19}, -\frac{16}{19})$

As above Corollaries 3.6 and 3.8 tell us that there are exactly two full-dimensional Voronoi cells of the curve in the simlex Lastly we look at the distance given by the matrix $M_3 =$

$$\begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

and notice that only the side along the vector (1,0,-1) gives tangency with the curve in the simplex, see Figure 2.2 for illustration. Hence there is exactly one full-dimensional Voronoi cell.

We will show that the Hardy-Weinberg curve always has one to three full-dimensional Voronoi cells w.r.t. a Wasserstein distance.

Lemma 4.1. *A Voronoi diagram of the Hardy-Weinberg curve in the simplex with Wasserstein distance has one, two or three full-dimensional Voronoi cells.*

Proof. We know from Corollary 3.6 that full-dimensional Voronoi cells can only appear where the tangent line of a point on the curve is the same as the line spanned by a facet of the Wasserstein distance ball. By the characterisation of the tangents we see that all points on the curve in the simplex have distinct tangent lines. This together with Proposition 2.5 and the fact that opposite facets are parallel tells us that we can have at most three full-dimensional cells. To

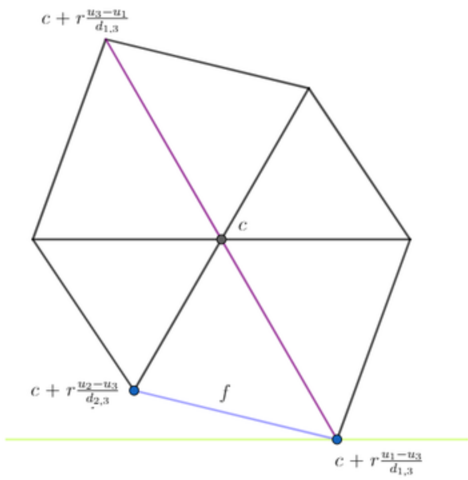

Figure 4.3: Illustration of proof of Lemma 4.1

argue that we have at least one full-dimensional cell we recall that the slope of the tangent lines of the Hardy-Weinberg curve in the simplex interpolated between the slopes of two sides of the simplex, one parallel to the purple segment in Figure 4.3. The third side of the simplex, parallel to the tangent line of the curve for $p = \frac{1}{2}$, is parallel to the light green line in Figure 4.3. Consider the two entries $d_{1,3}$ and $d_{2,3}$ from the Wasserstein distance matrix. By symmetry we may assume $d_{1,3} \leq d_{2,3}$. Now let f be the facet connecting $c + r\frac{u_1-u_3}{d_{1,3}}$ and $c + r\frac{u_2-u_3}{d_{2,3}}$. Since we assumed $d_{1,3} \leq d_{2,3}$, the slope of f is somewhere between that of a line through $c + u_1 - u_3$ and $c + u_2 - u_3$, parallel to light green line in Figure 4.3, and the diagonal through $c + r\frac{u_1-u_3}{d_{1,3}}$ and $c + r\frac{u_3-u_1}{d_{1,3}}$, which is parallel to one of the sides of the simplex. This puts the slope of f in an interval which defines some tangent line of the curve. Since the curve is convex this means we always have at least one full-dimensional Voronoi cell in the simplex by Corollary 3.8.

□

Chapter 5

Counting cells

In this section we find an upper bound for the number of full-dimensional Voronoi cells for smooth irreducible, varieties with hypersurfaces as dual varieties. These are a commonly studied objects which the Veronese curves we have discussed are examples of. We use dual varieties to find an upper bound on the number of points where any given facet of a Wasserstein ball can be tangent to the variety. We then present a separate but connected approach in the language of polar varieties and polar degrees that has been used in studying optimisation problems for Wasserstein distances in *Wasserstein distance to independence models* [3]. For reference of the theory on dual varieties used in this chapter see *Discriminants, Resultants and Multidimensional Determinants* [8].

So far we have studied Voronoi diagrams of manifolds in affine spaces. However to apply all the theory in this chapter we need to work with closed irreducible varieties in complex projective space. We will however be able to draw conclusion about our real affine variety $X \subset \mathbf{1}_n$ from what we learn about its complex projective closure $\overline{X}$.

5.1 Dual variety

We can find a projective space P^* for a projective space P by associating hyperplanes in P with points in the projective space P^*, and a point in P with the hyperplane in P^* of all hyperplanes in through the given point. This in fact gives us a duality between the two spaces where P^* is known as the projective dual space of P.

The closed irreducible projective variety $\overline{X} \subset P$ can now be said to have a projective dual variety $\overline{X}^\vee \in P^*$, the Zariski closure of the set of all hyperplanes tangent to $\overline{X}$ at some smooth point.

Fact 5.1. *Let $\overline{X} \subset P$ be a closed irreducible projective variety with dual variety $\overline{X}^{\vee}$ a hypersurface. For a generic projective subspace $V \subset P$ of codimension 2, the number of hyperplanes containing V and being tangent to $\overline{X}$ is the degree of the dual variety $\overline{X}^{\vee}$.*

5.2 Polar varieties and Polar degrees

We now introduce the concept of the i'th polar variety of a projective variety and projective subspaces. We will be using the same definition as in *Coisotropic Hypersurfaces in Grassmanians*[1].

Definition 5.2. *If in projective $n-$space we are given a smooth projective variety $\overline{X}$ of dimension k and a projective subspace of dimension n+i-k-2, called V, then the i'th polar variety of $\overline{X}$ with respect to V is defined as*

$$P_i(\overline{X}, V) = \overline{\{x \in \overline{X} : dim(T_x(\overline{X}) \cap V) \geq i-1\}}. \tag{5.1}$$

We are told in a discussion in *The geometric and numerical properties of duality in projective algebraic geometry*[2] that for any $\overline{X}$ there is an integer $\delta_i(\overline{X})$ which is the degree of $P_i(\overline{X}, V)$ for almost all V. These polar degrees $P_i(\overline{X}, V)$ have the following useful properties:

- The polar degree $\delta_i(\overline{X})$ is positive exactly when $k - codim(\overline{X}^{\vee}) + 1 \geq i \geq 0$.
- The 0'th polar degree $\delta_0(\overline{X})$ is the same as the degree of $\overline{X}$.
- The last nonzero polar degree $\delta_{k-codim(\overline{X}^{\vee})+1}(\overline{X})$ is the same as the degree of the dual variety $\overline{X}^{\vee}$.

To present the upper bound on full dimensional Voronoi cells promised in this chapter we introduce the notation $F(D)$ for the number of facets of the Wasserstein distance ball B_D

Theorem 5.3. *Given a smooth irreducible variety, $X \subset \mathbf{1}_n$ of dimension k such that the dual variety of its projective closure, $\overline{X}^{\vee}$, is a hypersurface, we have that for almost all Wasserstein distances D, the number of full-dimensional Voronoi cells of X is at most*

$$\frac{F(D)deg(\overline{X}^{\vee})}{2} = \frac{F(D)\delta_k(\overline{X})}{2}.$$

Proof. By Corollary 3.7 we know that only when the subspace spanned by a facet of the Wasserstein ball contains the tangent space of X at some point x can the Voronoi cell of x be full dimensional. By finding, for any hyperplane H, an upper bound M of the number of points on X which have their tangent space contained in a parallel translate of H we can give an upper bound for the number of full-dimensional cells. Every facet can then give at most M fulldimensional cells. Since opposite facets of the ball share the same slope and any point has only one Voronoi cell we get an upper bound of half the number of facets times M. Since parallel affine hyperplanes correspond to projective hyperplanes containing a common codimension 2 subspace $V \subset P$ and we assume a generic Wasserstein distance D, Fact 5.1 tells us that that $\deg(\overline{X}^{\vee}) = \delta_k(\overline{X})$ will work as our upper bound M. □

We note that the reasons the quantity in Theorem 5.3 is an upper bound are that not all points where a facet is tangent to the variety produce full-dimensional Voronoi cells in general and that the step from X to its complex projective closure can introduce additional tangency points.

Remark 5.4. *More generally, a face f of a generic distance ball B_D can also produce $(f+1)$-dimensional Voronoi cells and Theorem 3.5 shows that non-transversal intersection is needed for this to happen. This means that the dimension of the space spanned by the tangent space of X at some point x and a parallel translate of the span of f is less than the dimension n of the ambient space. Comparing this with the definition of the polar varieties (5.1), we see that $i = k - n + dim(f) + 1$ is sufficent. More precisely, the parallel translates of $span(f)$ in the affine ambient space correspond to projective subspaces of dimension dim(f) that contain a common subspace V of dimension dim$(f) - 1$ that lies in the hyperplane at infinity. Hence, if non-transversal intersection happens at a point $x \in X \subset \overline{X}$, then x must be in $P_i(\overline{X}, V)$. In other words, the points $x \in X$ such that the cell-cone component of the Voronoi cell $V_{D,X}(x)$ corresponding to f is not empty are all contained in the i'th polar variety of $\overline{X}$ with respect to V.*

We expect infinitely many $(f+1)$-dimensional Voronoi cells when f is not a facet, so the polar degree $\delta_i(\overline{X})$ will not give the number of these Voronoi cells. However, this polar degree is an upper bound for the number of critical points of an optimization problem restricted to a face of the Wasserstein ball in Theorem 13 and Proposition 17 of Wasserstein distance to independence models[3]. Figure 3.2 displays a critical point of this kind.

It has been proved [6] that in a $n-$dimensional metric space, the number of $(m-1)-$dimensional faces of a generic distance ball is

$$\frac{(n+m)!}{(m!)^2(n-m)!}.$$

This result tells us the expected number of facets of a generic Wasserstein ball. This expression is also a maximum for facets[7] where the expression becomes

$$\binom{2n}{n}.$$

We now give an example of when the bound in Theorem 5.3 is sharp.

Example 5.5. *Consider the $n-$sphere in an $(n+1)-$dimensional ambient space. The $n-$sphere has another $n-$sphere as its dual variety which has degree 2. This means that the upper bound of full-dimensional Voronoi cells is the number of facets of the Wasserstein ball B_D. We also know that for every hyperplane there are exactly 2 parallel translates that are tangent to the $n-$sphere. If a hyperplane passes through the point x on the $n-$sphere where it is tangent, then one of the two closed halfspaces created intersects the $n-$sphere only at x. Therefore we can conclude that every facet of B_D contributes with one full-dimensional Voronoi cell. This is the same number of full-dimensional cells as predicted by the upper bound.*

Bibliography

[1] Kathlén Kohn. *Coisotropic Hypersurfaces in Grassmannians Journal of Symbolic Computation 103 (2021) 157-177.*

[2] Holme, Audun.
The geometric and numerical properties of duality in projective algebraic geometry
Manuscripta mathematica 61.2 (1988): 145-162.

[3] Türkü Özlüm Celik, Asgar Jamneshan, Guido Montúfar, Bernd Sturmfels, Lorenzo Venturello *Wasserstein distance to independence models Journal of Symbolic Computation 104 (2021) 855-873.*

[4] Atsuyuki Okabe, Barry Boots, Kokichi Sugihara, Sung Nok Chiu *Spacial Tessellations concepts and applications of Voronoi diagrams 2'ed.*
John Wiles & sons Ltd, 2000.

[5] Diego Cifuentes, Kristian Ranestad, Bernd Sturmfels, Madeleine Weinstein *Voronoi cells of varieties*
Journal of Symbolic Computation (2020)..

[6] J. Gordon, F. Petrov *Combinatorics of the Lipschitz polytope Arnold Mathematical Journal 3.2 (2017): 205-218..*

[7] Michael Joswig, Katja Kulas *Tropical and ordinary convexity combined Walter de Gruyter GmbH & Co. KG 10.2 (2010), 333-352.*

[8] I. M. Gelfand, M. M. Kapranov, A. V. Zelevinsky *Discriminants, resultants, and multidimensional determinants*
Springer Science & Business Media, 2008.

[9] Rüschendorf, L. (2001) [1994], *Wasserstein metric*, Encyclopedia of Mathematics, EMS Press
URL: `http://encyclopediaofmath.org/index.php?title=Wasserstein_metric&oldid=50083`

TRITA-SCI-GRU 2021:214