

# Metric Algebraic Geometry

Maddie Weinstein

University of Puget Sound

*mweinstein@pugetsound.edu*

# Overview

- 1 Background on Algebraic Varieties and Metric Algebraic Geometry
- 2 Metric Algebraic Geometry and Data
  - Learning Algebraic Varieties from Samples
  - An Algebro-Geometric Perspective on Topological Data Analysis
  - Reach: Function of Bottlenecks and Curvature
  - Bottlenecks
  - Curvature
- 3 Metric Algebraic Geometry and Optimization
  - Voronoi Cells: From Point Sets to Positive Dimension
  - Application: Low-Rank Matrix Approximation

## Background on Algebraic Varieties and Metric Algebraic Geometry

# Varieties and Algebraic Geometry



**Figure:** A curve and a surface in  $\mathbb{R}^3$ . The curve is the zero set of the polynomials  $x_1^2 + x_2^2 + x_3^2 - 4$  and  $(x_1 - 1)^2 + x_2^2 - 1$ . The surface is defined by a degree 10 polynomial.

An **algebraic variety** (a.k.a. variety) is the zero set of a system of polynomial equations. **Algebraic geometry** is the study of the geometry of algebraic varieties and their generalizations.

# Degree of an Algebraic Variety

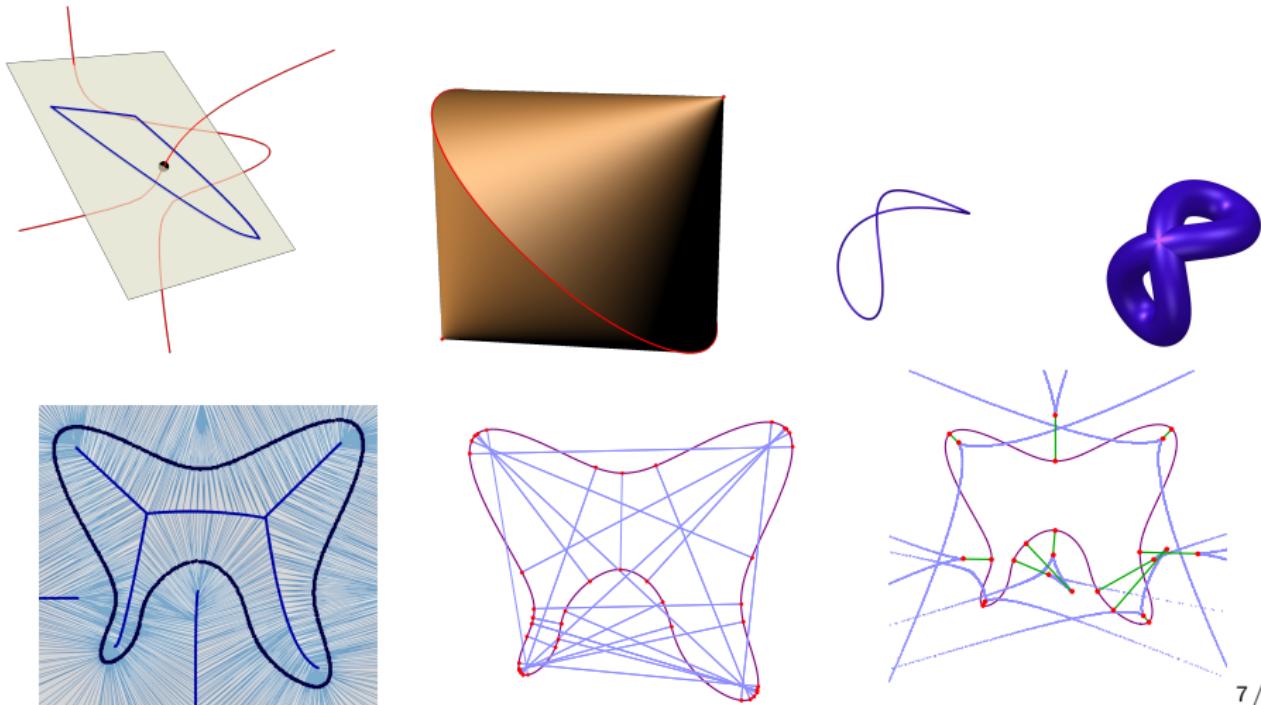
- The degree of a variety defined by one polynomial is the degree of the polynomial.
- For varieties defined by more than one polynomial, it is harder to determine the degree. The field of **intersection theory** studies these questions.
- Degree is a proxy for computational complexity.
- In  $\mathbb{R}^n$ , a variety of degree  $d$  intersects a linear space of complementary dimension in at most  $d$  points. For a consistent count, we turn to **complex projective space**.

# Projective Space

- We switch from  $\mathbb{R}^n$  to  $\mathbb{C}^n$  when we want every polynomial of degree  $d$  to have  $d$  roots.
- We switch from **affine space** to **projective space** when we want every pair of lines to intersect.
- Example:  $\mathbb{P}_{\mathbb{R}}^2 = \mathbb{R}^2 \cup \{\text{one point at } \infty \text{ for each equivalence class of parallel lines}\}$
- *For a more thorough introduction for students, I recommend the textbook **Ideals, Varieties, and Algorithms** by Cox, Little, and O'Shea.*

# What is metric algebraic geometry?

**Metric algebraic geometry** is concerned with properties of **real algebraic varieties** that depend on a **metric**. Results can be applied to **distance optimization**, **algebraic statistics**, and the study of the **geometry of data** with **nonlinear models**.



# Metric Algebraic Geometry and Data

# Application: Computational Study of Geometry of Data with Nonlinear Models

Suppose we are given of sample of points in  $\mathbb{R}^n$  and we choose to model them with an algebraic variety  $V$ . What can we learn about the variety?

In **Learning Algebraic Varieties from Samples**, joint with Paul Breiding, Sara Kalisnik, Bernd Sturmfels, we survey, develop, and implement methods to study the following properties:

- What is the **dimension** of  $V$ ?
- What **equations** vanish on  $V$ ?
- What is the **homology** of  $V$ ?

# Topological Data Analysis

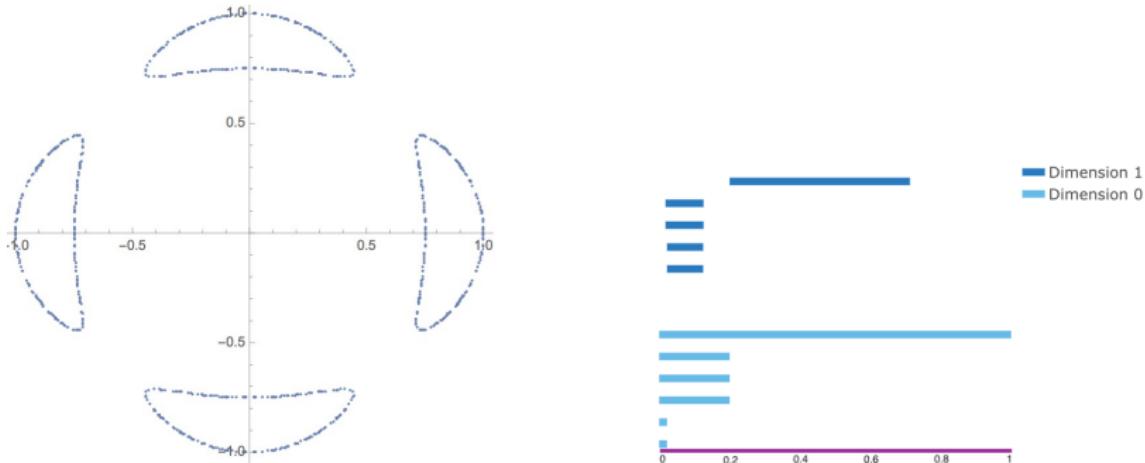


Figure: Persistent homology barcodes for the Trott curve in  $\mathbb{R}^2$ .

- We draw a ball of radius  $r$  around each point and count the number of “holes” in the resulting union of balls.
- The number of light blue bars at a given radius is the number of connected components of the union of balls of that radius.
- The number of dark blue bars at a given radius is the number of two dimensional holes in the union of balls of that radius.

# Algebraicity of Persistent Homology

In **Offset Hypersurfaces and Persistent Homology of Algebraic Varieties**, joint with Emil Horobet, we study the persistent homology of the **offset filtration of an algebraic variety**, bringing the perspective of real algebraic geometry to the study of persistent homology and showing that barcodes can be computed exactly.



Figure: The Viviani curve (left) and its  $\epsilon = 1$  offset surface (right).

## Theorem (Horobet-W.)

Let  $f_1, \dots, f_s$  be polynomials in  $\mathbb{Q}[x_1, \dots, x_n]$  with  $X_{\mathbb{R}} = V_{\mathbb{R}}(f_1, \dots, f_s)$ . Then the values of the persistence parameter  $\epsilon$  at which a bar in the offset filtration barcode appears or disappears are real numbers algebraic over  $\mathbb{Q}$ .

# Guaranteeing Persistent Homology: The Reach of an Algebraic Variety

Theorem (Partha Niyogi, Stephen Smale, Shmuel Weinberger 2006)

Let  $M$  be a compact submanifold of  $\mathbb{R}^N$  of dimension  $k$  with **reach**  $\tau$ . Let  $\bar{x} = \{x_1, \dots, x_n\}$  be a set of  $n$  points drawn in independent and identically distributed fashion according to the uniform probability measure on  $M$ . Let  $0 < \epsilon < \frac{\tau}{2}$ . Let

$U = \bigcup_{x \in \bar{x}} B_\epsilon(x)$  be a corresponding random open subset of  $\mathbb{R}^N$ . Let  $\beta_1 = \frac{\text{vol}(M)}{(\cos^k(\theta_1))\text{vol}(B_\epsilon^k/4)}$  and  $\beta_2 = \frac{\text{vol}(M)}{(\cos^k(\theta_2))\text{vol}(B_\epsilon^k/8)}$ , where  $\theta_1 = \arcsin(\frac{\epsilon\tau}{8})$  and  $\theta_2 = \arcsin(\frac{\epsilon\tau}{16})$ . Then for all

$$n > \beta_1 \left( \log(\beta_2) + \log\left(\frac{1}{\delta}\right) \right)$$

the homology of  $U$  equals the homology of  $M$  with high confidence (probability  $> 1 - \delta$ ).

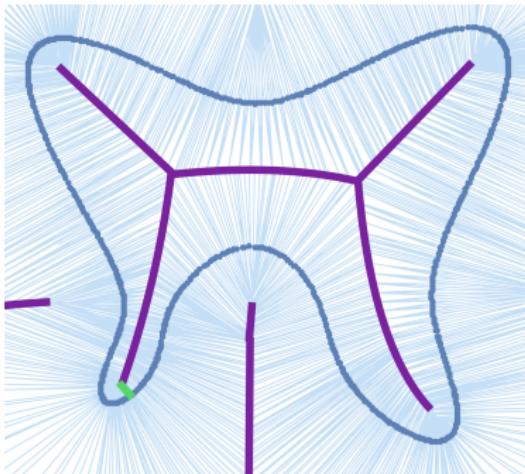
## Highlight

The **reach** (to be defined on the next slide) of an algebraic variety determines the number of sample points required for persistent homology to work.

# The Reach of an Algebraic Variety

## Definition

The **medial axis** of a variety  $V \subset \mathbb{R}^n$  is the set of all points  $u \in \mathbb{R}^n$  such that the minimum distance from  $V$  to  $u$  is attained by two distinct points. The **reach**  $\tau_V$  is the infimum of all distances from points on the variety  $V$  to points in its medial axis.



**Figure:** The medial axis (purple) and reach (green) of the quartic butterfly curve can be seen in its Voronoi approximation.

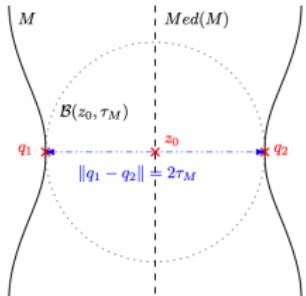
# Algebraicity of Reach

## Proposition (Horobet-W.)

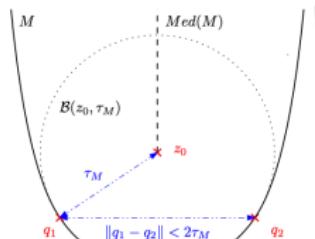
*Let  $V$  be a smooth algebraic variety in  $\mathbb{R}^n$ . Let  $f_1, \dots, f_s \in \mathbb{Q}[x_1, \dots, x_n]$  with  $V = V_{\mathbb{R}}(f_1, \dots, f_s)$ . Then the reach of  $V$  is an algebraic number over  $\mathbb{Q}$ .*

This theorem is an invitation to use algebraic geometry to find the reach of a variety.

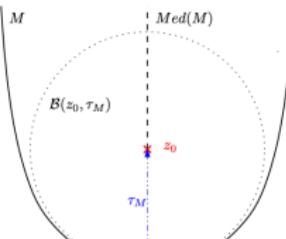
# Reach, Bottlenecks, and Curvature



(a) A bottleneck.



(b) A non-bottleneck reach attaining pair.



(c) Curvature-attaining point.

**Figure:** The reach of a manifold is attained by a bottleneck, two points on a circular arc, or a point of maximal curvature. Figure and Theorem due to Aamari-Kim-Chazal-Michel-Rinaldo-Wasserman '17.

# Bottlenecks

## Definition

Let  $V \subset \mathbb{R}^n$  be a smooth variety. A line is **orthogonal** to  $V$  if it is orthogonal to the tangent space  $T_x V \subset \mathbb{R}^n$  at  $x$ . The **bottlenecks** of  $V$  are pairs  $(x, y)$  of distinct points  $x, y \in V$  such that the line spanned by  $x$  and  $y$  is orthogonal to  $V$  at both points.

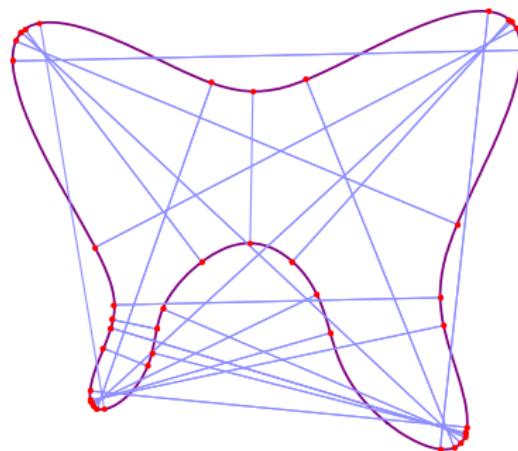


Figure: The real bottleneck pairs of the quartic butterfly curve.

# Bottlenecks as Distance Optimization

## Remark

*Bottlenecks are the critical points of the squared distance function*

$$\mathbb{R}^n \times \mathbb{R}^n : (x, y) \mapsto \|x - y\|^2,$$

*subject to the constraints  $x, y \in V$  as well as the non-triviality condition  $x \neq y$ .*

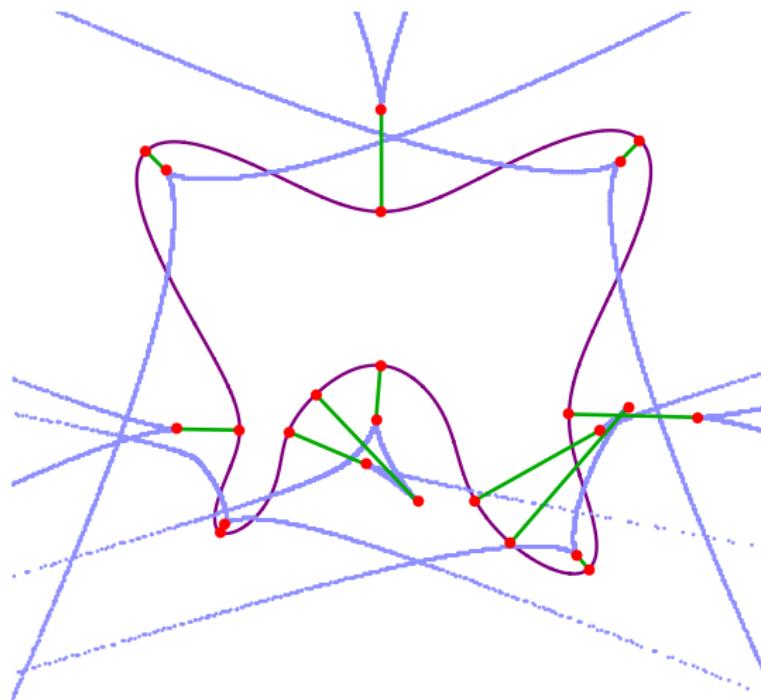
# Bottleneck Degree

Denote by  $BND(V)$  the bottleneck degree of  $V \subset \mathbb{C}^n$ . Under certain conditions, this coincides with twice the number of bottleneck pairs.

## Theorem (Di Rocco-Eklund-W.)

- Let  $V \subset \mathbb{C}^2$  be a “general” curve of degree  $d$ . Then  
$$BND(V) = d^4 - 5d^2 + 4d.$$
- Let  $V \subset \mathbb{C}^3$  be a “general” surface of degree  $d$ . Then  
$$BND(V) = d^6 - 2d^5 + 3d^4 - 15d^3 + 26d^2 - 13d.$$
- For any smooth variety  $V \subset \mathbb{P}_{\mathbb{C}}^n$  in “general position,” we have an algorithm to express the bottleneck degree in terms of the polar classes of  $V$ .

# Curvature



**Figure:** The real points of critical curvature on the butterfly curve (purple) joined by green line segments to their centers of curvature. These give cusps on the evolute (light blue).

# Degree of Critical Curvature

## Theorem (Salmon, 1852)

Let  $V \subset \mathbb{R}^2$  be a general algebraic curve of degree  $d$ . Then the number of complex critical curvature points of  $V$  counted with multiplicity is  $6d^2 - 10d$ .

## Theorem (Breiding-Ranestad-W.)

Let  $V \subset \mathbb{R}^3$  be a general algebraic surface of degree  $d$ . Then  $V$  has isolated complex critical curvature points. An upper bound for their number is  $\frac{1}{8}(2796d^3 - 6444d^2 + 3696d)$ .

A formula for hypersurfaces in  $\mathbb{R}^n$  can be computed using the same process given in our proof. However, we do not have a proof that a general hypersurface in  $\mathbb{R}^n$  for  $n \geq 4$  has isolated complex critical curvature points.

# Metric Algebraic Geometry and Optimization

# Voronoi Diagrams of Point Clouds

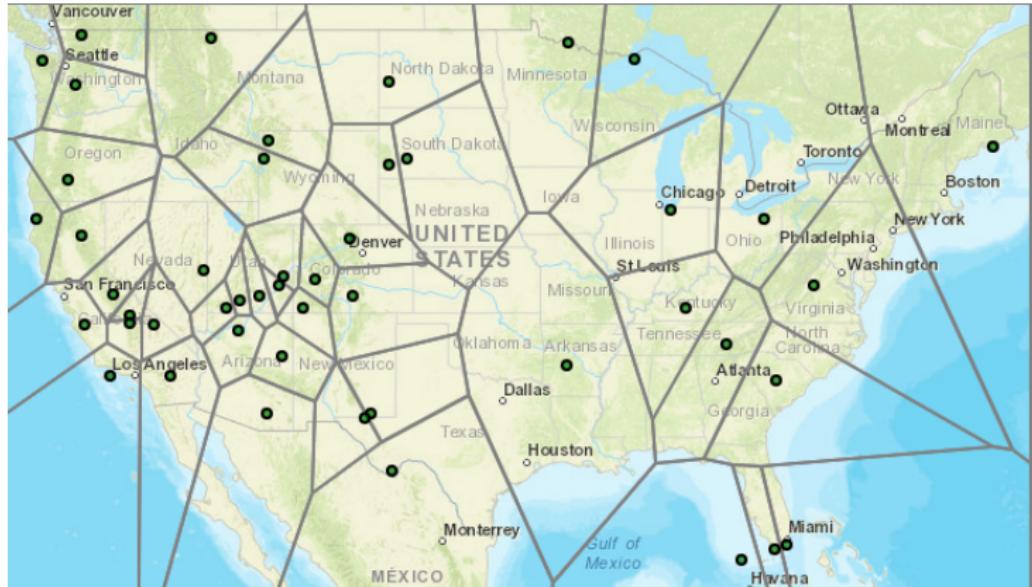


Figure: A Voronoi diagram of national parks. Source:  
<https://engaging-data.com/national-park-voronoi/>

# Voronoi Cells of Varieties

## Definition

Let  $V$  be a real algebraic variety of codimension  $c$  in  $\mathbb{R}^n$  and  $y$  a smooth point on  $V$ . Its *Voronoi cell* consists of all points whose closest point in  $V$  is  $y$ , i.e.

$$\text{Vor}_V(y) := \left\{ u \in \mathbb{R}^n : y \in \arg \min_{x \in V} \|x - u\|^2 \right\}.$$

The Voronoi cell  $\text{Vor}_V(y)$  is a convex semialgebraic set of dimension  $c$ , living in the normal space  $N_V(y)$  to  $V$  at  $y$ . Its boundary consists of the points in  $\mathbb{R}^n$  that have at least two closest points in  $V$ , including  $y$ .

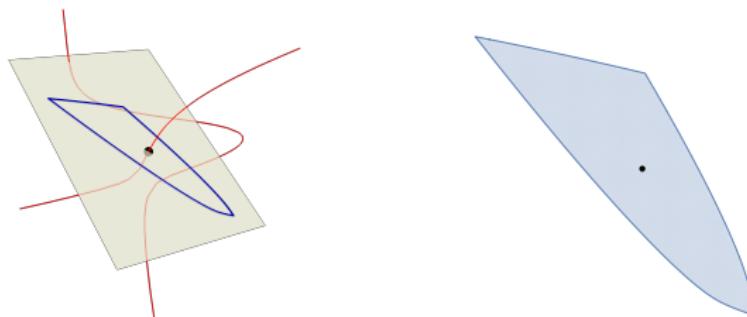
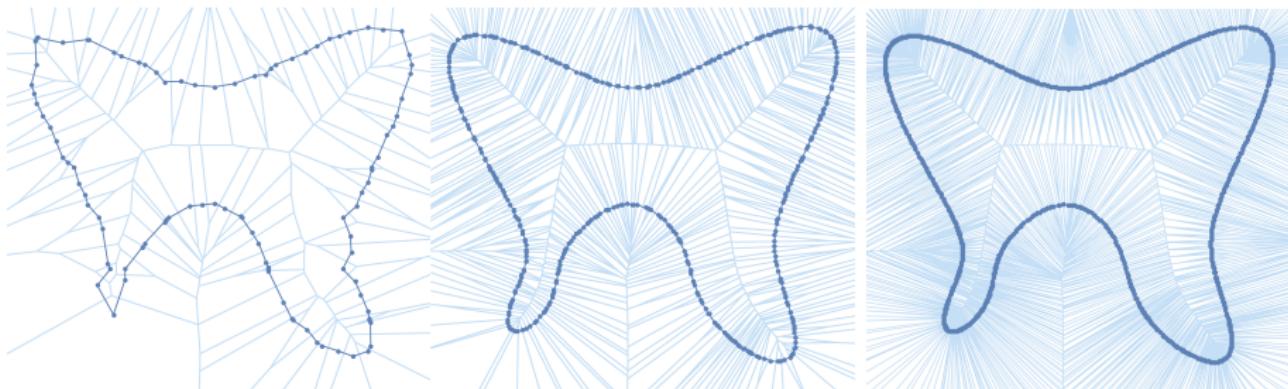


Figure: A quartic space curve and the Voronoi cell in one of its normal planes.

# Voronoi Convergence

## Theorem (Brandt-W.)

*As sampling density increases, the Voronoi diagrams of a point sample of a variety “converge” to the Voronoi decomposition of the variety.*



**Figure:** Voronoi cells of 101, 441, and 1179 points sampled from the quartic butterfly curve:  $x^4 - x^2y^2 + y^4 - 4x^2 - 2y^2 - x - 4y + 1 = 0$ .

# Voronoi Degrees

## Remark

When discussing degree, we identify the variety  $V$  and  $\delta_{\text{alg}} \text{Vor}_V(y)$  with their Zariski closures in  $\mathbb{P}_{\mathbb{C}}^n$ .

## Definition

The algebraic boundary of the Voronoi cell  $\text{Vor}_V(y)$  is a hypersurface in the normal space to  $V$  at  $y$ . Its degree  $\delta_V(y)$  is called the *Voronoi degree* of  $V$  at  $y$ .

# Voronoi Degrees

## Theorem (Cifuentes-Ranestad-Sturmfels-W.)

- Let  $V \subset \mathbb{P}^n$  be a curve of degree  $d$  and geometric genus  $g$  with at most ordinary multiple points as singularities. The Voronoi degree at a general point  $y \in V$  equals

$$\delta_V(y) = 4d + 2g - 6,$$

provided  $V$  is in general position in  $\mathbb{P}^n$ .

- Let  $V \subset \mathbb{P}^n$  be a smooth surface of degree  $d$ . Then its Voronoi degree equals

$$\delta_V(y) = 3d + \chi(V) + 4g(V) - 11,$$

provided the surface  $V$  is in general position in  $\mathbb{P}^n$  and  $y$  is a general point on  $V$ , where  $\chi(V) := c_2(V)$  is the topological Euler characteristic and  $g(V)$  is the genus of the curve obtained by intersecting  $V$  with a general smooth quadratic hypersurface in  $\mathbb{P}^n$ .

## Application: Low-Rank Matrix Approximation

The study of Voronoi cells can be applied to low-rank matrix approximation.

# Norms on Space of Matrices

Fix the space  $\mathbb{R}^{m \times n}$  of real  $m \times n$  matrices.

## Frobenius Norm

$$\|U\|_F := \sqrt{\sum_{ij} U_{ij}^2}$$

## Spectral Norm

$$\|U\|_2 := \max_i \sigma_i(U) \text{ which extracts the largest singular value.}$$

# Low-Rank Matrix Approximation

## Remark

Let  $X$  denote the variety in  $\mathbb{R}^{m \times n}$  of real  $m \times n$  matrices of rank  $\leq r$ . Fix a rank  $r$  matrix  $V$  in  $X$ . Let  $U \in \text{Vor}_X(V)$  and let  $U = \Sigma_1 D \Sigma_2$  be its singular value decomposition. Let  $D^{[r]}$  be the matrix that is obtained from  $D$  by replacing all singular values except for the  $r$  largest ones by zero. By the Eckart-Young Theorem, we have  $V = \Sigma_1 \cdot D^{[r]} \cdot \Sigma_2$ .

## Remark

The Eckart-Young Theorem works for both norms, so both give the same Voronoi cell  $\text{Vor}_X(V)$ .

# Norms on Space of Symmetric Matrices

Consider the space  $\mathbb{R}^{\binom{n+1}{2}}$  whose coordinates are the upper triangular entries of a symmetric  $n \times n$  matrix. Let  $X$  be the variety of symmetric matrices of rank  $\leq r$ .

## Remark

*The Frobenius norm and Euclidean norm differ on this space.*

The Frobenius norm on  $\mathbb{R}^{\binom{n+1}{2}}$  is the restriction of the Frobenius norm on  $\mathbb{R}^{n \times n}$  to the subspace of symmetric matrices.

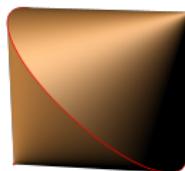
## Example

Let  $n = 2$ . We identify the vector  $(a, b, c)$  with the symmetric matrix  $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$ . The Frobenius norm is  $\sqrt{a^2 + 2b^2 + c^2}$ , whereas the Euclidean norm is  $\sqrt{a^2 + b^2 + c^2}$ .

# Low-Rank Approximation on the Space of Symmetric Matrices

## Remark

*The Frobenius and Euclidean norms have dramatically different properties with respect to low-rank approximation of symmetric matrices. The Eckart-Young Theorem is valid for the Frobenius norm on  $\mathbb{R}^{\binom{n+1}{2}}$  but not for the Euclidean norm.*



**Figure:** The Voronoi cell of a symmetric  $3 \times 3$  matrix of rank 1 is a convex body of dimension 3. It is shown for the Frobenius norm (left) and for the Euclidean norm (right).

Thank you!

## Polar Classes

An  $m$ -dimensional variety has  $m + 1$  polar varieties defined by exceptional tangent loci.

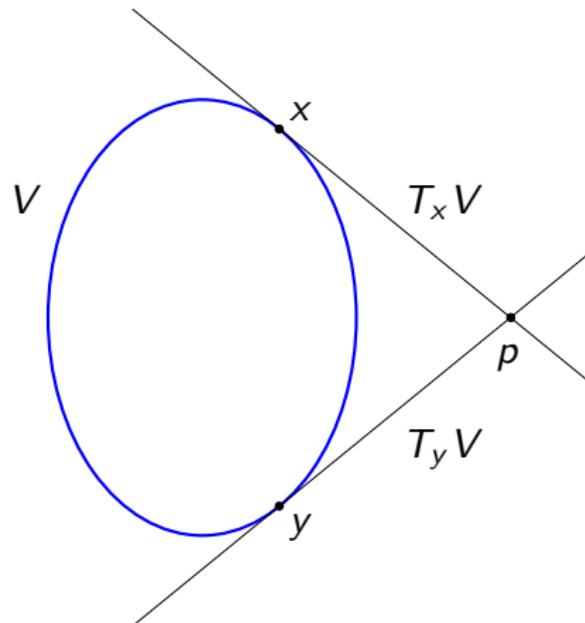


Figure: Polar locus of a point  $p$  with respect to an ellipse  $V$ :  $P_1(V, p) = \{x, y\}$ .

# Polar Classes

## Example

For a smooth surface  $V \subset \mathbb{P}^3$  we have two polar varieties. Let  $p \in \mathbb{P}^3$  be a general point and  $I \subset \mathbb{P}^3$  a general line. Then  $P_1(V, p)$  is the set of points  $x \in V$  such that the projective tangent plane  $\mathbb{T}_x V \subset \mathbb{P}^3$  contains  $p$ . This is a curve on  $V$ . Similarly,  $P_2(V, I) = \{x \in V : I \subseteq \mathbb{T}_x V\}$ , which is finite.

## Definition

Let  $V \subset \mathbb{P}^n$  be a smooth variety of dimension  $m$ . For  $j = 0, \dots, m$  and a general linear space  $L \subseteq \mathbb{P}^n$  of dimension  $n - m - 2 + j$  the **polar variety** is given by

$$P_j(V, L) = \{x \in V : \dim \mathbb{T}_x V \cap L \geq j - 1\}.$$

For each polar variety  $P_j(V, L)$ , there is a corresponding **polar class**  $[P_j(V, L)] = p_j$  which represents  $P_j(V, L)$  up to rational equivalence.

## Polar Classes and Chern Classes

$P_j(V, L)$  is either empty or of pure codimension  $j$  and

$$p_j = \sum_{i=0}^j (-1)^i \binom{m-i+1}{j-i} h^{j-i} c_i(T_X),$$

where  $h \in A_{n-1}(X)$  is the hyperplane class.

The polar loci  $P_j(V, L)$  are reduced. We have

$$c_j(T_X) = \sum_{i=0}^j (-1)^i \binom{m-i+1}{j-i} h^{j-i} p_i.$$

## Bottleneck Genericity Assumptions 1/2

Let  $V \subset \mathbb{P}^n$  be a variety. Consider the **conormal variety**

$$\mathcal{C}_V = \{(p, q) \in \mathbb{P}^n \times \mathbb{P}^n : p \in V, q \in (\mathbb{T}_p V)^\perp\}$$

and the map

$$f : \mathcal{C}_V \rightarrow \text{Gr}(2, n+1) : (p, q) \mapsto \langle p, q \rangle$$

from  $\mathcal{C}_V$  to the Grassmannian of lines in  $\mathbb{P}^n$  that sends a pair  $(p, q)$  to the line spanned by  $p$  and  $q$ .

The orthogonality relation on  $\mathbb{P}^n$  is defined via the **isotropic quadric**  $Q \subset \mathbb{P}^n$  given in homogeneous coordinates by  $\sum_0^n x_i^2 = 0$ . Varieties which are tangent to  $Q$  are to be considered degenerate in this context and we say that a smooth projective variety is in **general position** if it intersects  $Q$  transversely. Equivalently, a smooth variety  $V \subset \mathbb{P}^n$  is in general position if  $\mathcal{C}_V$  is disjoint from the diagonal  $\Delta \subset \mathbb{P}^n \times \mathbb{P}^n$ .

## Bottleneck Genericity Assumptions 2/2

A smooth variety  $V \subset \mathbb{P}^n$  is **bottleneck regular** if

- ①  $V$  is in general position,
- ②  $V$  has only finitely many bottlenecks and
- ③ the differential  $df_p : T_p \mathcal{C}_V \rightarrow T_{f(p)} G$  of the map  $f$  has full rank for all  $p \in \mathcal{C}_V$ .

If  $V \subset \mathbb{P}^n$  is bottleneck regular, then  $V$  is equal to the number of bottlenecks of  $V$  counted with multiplicity.

# Voronoi Genericity Assumptions

Let  $V \subset \mathbb{P}^n$  be a smooth projective variety defined over  $\mathbb{R}$ . We assume that  $y \in V$  is a general point, and that we fixed an affine space  $\mathbb{R}^n \subset \mathbb{P}^n$  containing  $y$  such that the hyperplane at infinity  $\mathbb{P}^n \setminus \mathbb{R}^n$  is in general position with respect to  $V$ .

## Example

For example, let  $V$  be the twisted cubic curve in  $\mathbb{P}^3$ , with affine parameterization  $t \mapsto (t, t^2, t^3)$ . Here  $g = 0$  and  $d = 3$ , so the expected Voronoi degree is 6.

However, for  $V \subset \mathbb{R}^3$  defined by the two equations  $\mathbf{f} = (x_1^2 - x_2, x_1 x_2 - x_3)$ , we can compute that algebraic boundary of  $\text{Vor}_V(0)$  is given by the quartic curve  $27u_3^4 + 128u_2^3 + 72u_2u_3^2 - 160u_2^2 - 35u_3^2 + 66u_2 = 9$  in the plane  $u_1=0$ . So  $\delta_V(y) = 4$ .

This is explained by the fact that the plane at infinity in  $\mathbb{P}^3$  intersects the curve  $V$  in a triple point. After a general linear change of coordinates in  $\mathbb{P}^3$ , which amounts to a linear fractional transformation in  $\mathbb{R}^3$ , we correctly find  $\delta_V(y) = 6$ .

# Convergence Theorem: Voronoi Version

## Definition

Given a point  $x \in \mathbb{R}^n$  and a closed set  $B \subset \mathbb{R}^n$ , define

$$d_w(x, B) = \inf_{b \in B} d(x, b).$$

A sequence  $\{B_\nu\}_{\nu \in \mathbb{N}}$  of compact sets is *Wijsman convergent* to  $B$  if for every  $x \in \mathbb{R}^n$ , we have that

$$d_w(x, B_\nu) \rightarrow d_w(x, B).$$

## Theorem (Brandt-W.)

Let  $X$  be a compact curve in  $\mathbb{R}^2$  and  $\{A_\epsilon\}_{\epsilon \searrow 0}$  a sequence of finite subsets of  $X$  containing all singular points of  $X$  such that every point of  $X$  is within distance  $\epsilon$  of some point in  $A_\epsilon$ . Then every Voronoi cell of  $X$  is the Wijsman limit of a sequence of Voronoi cells of  $\{A_\epsilon\}_{\epsilon \searrow 0}$ .

## Convergence Theorem: Delaunay Version

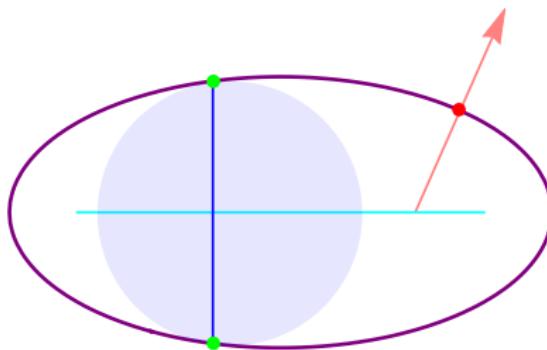
### Theorem (Brandt-W.)

Let  $X$  be a compact curve in  $\mathbb{R}^2$  and  $\{A_\epsilon\}_{\epsilon \searrow 0}$  a sequence of finite subsets of  $X$  containing all singular points of  $X$  such that no point of  $X$  is more than distance  $\epsilon$  from every point in  $A_\epsilon$ . If  $X$  is not tangent to any circle in four or more points, then every maximal Delaunay cell is the Hausdorff limit of a sequence of Delaunay cells of  $\{A_\epsilon\}_{\epsilon \searrow 0}$ .

# Delaunay Cells

## Definition

Let  $B(p, r)$  denote the open disc with center  $p \in \mathbb{R}^n$  and radius  $r > 0$ . We say this disc is *inscribed* with respect to  $X$  if  $X \cap B(p, r) = \emptyset$  and we say it is *maximal* if no disc containing  $B(p, r)$  shares this property. Given an inscribed disc  $B$  of an algebraic variety  $X \subset \mathbb{R}^n$ , the *Delaunay cell*  $\text{Del}_X(B)$  is  $\text{conv}(\overline{B} \cap X)$ .



**Figure:** The dark blue line segment is a Delaunay cell defined by the light blue maximally inscribed circle with center  $(-3/8, 0)$  and radius  $\sqrt{61}/8$ .

# Duality of Delaunay and Voronoi Cells

## Definition

Let  $X \subset \mathbb{R}^2$  be a finite point set. A *Delaunay triangulation* is a triangulation  $DT(X)$  of  $X$  such that no point of  $X$  is inside the circumcircle of any triangle of  $DT(X)$ .

## Remark

*The circumcenters of triangles in  $DT(X)$  are the vertices in the Voronoi diagram of  $X$ .*

# Hausdorff Convergence

The *Hausdorff distance* of two compact sets  $B_1$  and  $B_2$  in  $\mathbb{R}^n$  is defined as

$$d_h(B_1, B_2) := \sup \left\{ \sup_{x \in B_1} \inf_{y \in B_2} d(x, y), \sup_{y \in B_2} \inf_{x \in B_1} d(x, y) \right\}.$$

If an adversary gets to put your ice cream on either set  $B_1$  or  $B_2$  with the goal of making you go as far as possible, and you get to pick your starting place in the opposite set, then  $d_h(B_1, B_2)$  is the farthest the adversary could make you walk in order for you to reach your ice cream.

## Definition

A sequence  $\{B_\nu\}_{\nu \in \mathbb{N}}$  of compact sets is *Hausdorff convergent* to  $B$  if  $d_h(B, B_\nu) \rightarrow 0$  as  $\nu \rightarrow \infty$ .