

# Interpolation via Barycentric Coordinates

Pierre Alliez

Inria

Some material from D. Anisimov

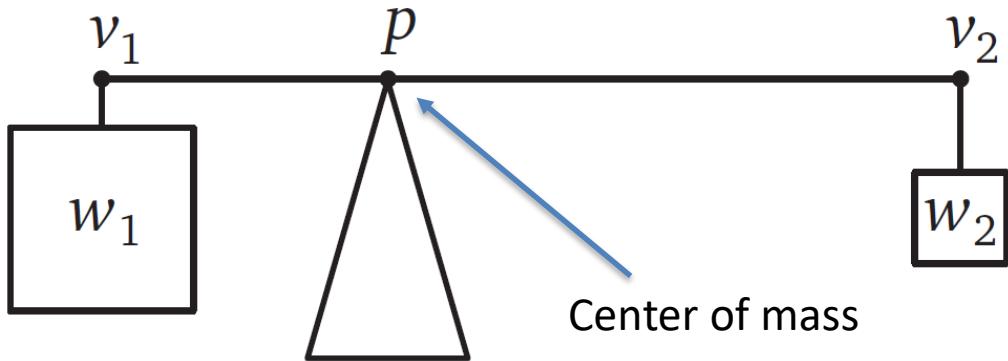
# Outline

- Barycenter
- Convexity
- Barycentric coordinates
  - For Simplices
  - For point sets
    - Inverse distance (Shepard)
    - Delaunay / Voronoi
    - Natural neighbors (Sibson)
  - For convex polyhedra
  - Generalized barycentric coordinates
  - Applications

**BARYCENTER**

# Barycenter

- Law of lever:  $w_1 l_1 = w_2 l_2$



$$l_1 = p - v_1$$

$$l_2 = v_2 - p$$

# Barycentric coordinates

$$w_1(v_1 - p) + w_2(v_2 - p) = 0$$

$$w_1 v_1 + w_2 v_2 = W p$$

$$b_1 v_1 + b_2 v_2 = p$$

$$w_1 + w_2$$

# Opposite problem?

$$b_1 v_1 + b_2 v_2 = p$$

?



# Opposite problem

$$b_1 v_1 + b_2 v_2 = p$$

$l_2/l$

$l_1/l$

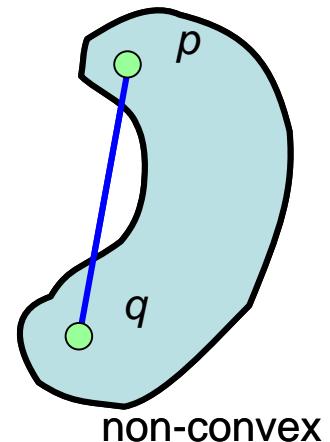
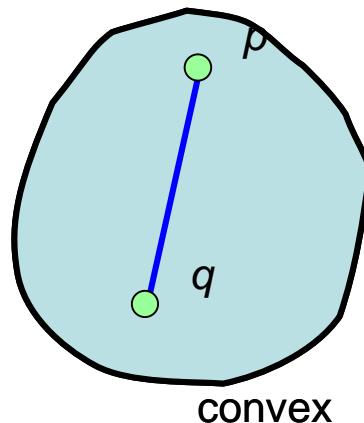
unique!



# **CONVEXITY**

# Convexity

A set  $S$  is *convex* if any pair of points  $p,q \in S$  satisfy  $pq \subseteq S$ .

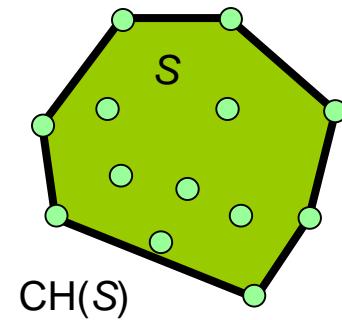
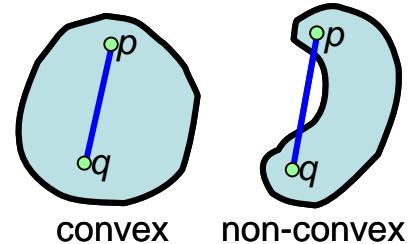


# Convex Hull

- The *convex hull* of a set  $S$  is:

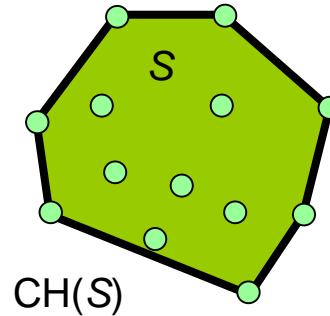
- The minimal convex set that contains  $S$ , i.e. any convex set  $C$  such that  $S \subseteq C$  satisfies  $\text{CH}(S) \subseteq C$ .
- The intersection of all convex sets that contain  $S$ .
- The set of all convex combinations of  $p_i \in S$ , i.e. all points of the form:

$$\sum_{i=1}^n \alpha_i p_i , \quad \alpha_i \geq 0, \quad \sum_{i=1}^n \alpha_i = 1$$



# Convex Hulls

- The convex hull of a set is unique, up to collinearities.
- The boundary of the convex hull of a point set is a polygon on a subset of the points.



# **BARYCENTRIC COORDINATES FOR SIMPLICES**

# Triangle barycentric coordinates

$$b_3 = \frac{A_3}{A}$$
$$b_1 = \frac{A_1}{A} \quad b_2 = \frac{A_2}{A}$$

$$A = A_1 + A_2 + A_3$$



A. F. Möbius

A. F. Möbius  
[1790–1868]

# Triangle barycentric coordinates

$$b_1 = \frac{A_1}{A} \quad A = A_1 + A_2 + A_3$$
$$b_2 = \frac{A_2}{A}$$
$$b_3 = \frac{A_3}{A}$$

The diagram shows a triangle  $T$  with vertices  $v_1$ ,  $v_2$ , and  $v_3$ . The triangle is divided into three smaller triangles by dashed lines from vertex  $v$  to each side. The areas of these smaller triangles are labeled  $A_1$  (red),  $A_2$  (green), and  $A_3$  (blue). The total area  $A$  is the sum of  $A_1$ ,  $A_2$ , and  $A_3$ . The barycentric coordinates  $b_1$ ,  $b_2$ , and  $b_3$  are defined as the ratios of these areas to the total area  $A$ .

**Properties:**

Constant precision:  $b_1 + b_2 + b_3 = 1$

Linear precision:  $b_1 v_1 + b_2 v_2 + b_3 v_3 = v$

The Lagrange property:  $b_i(v_j) = \delta_{ij}$

Non-negativity:  $b_i \geq 0$  for  $i = 1, 2, 3$

Linearity along edges:  $b_i = av + b$  for  $v \in \partial T$

Smoothness:  $b_i \in C^k$ ,  $k > 0$

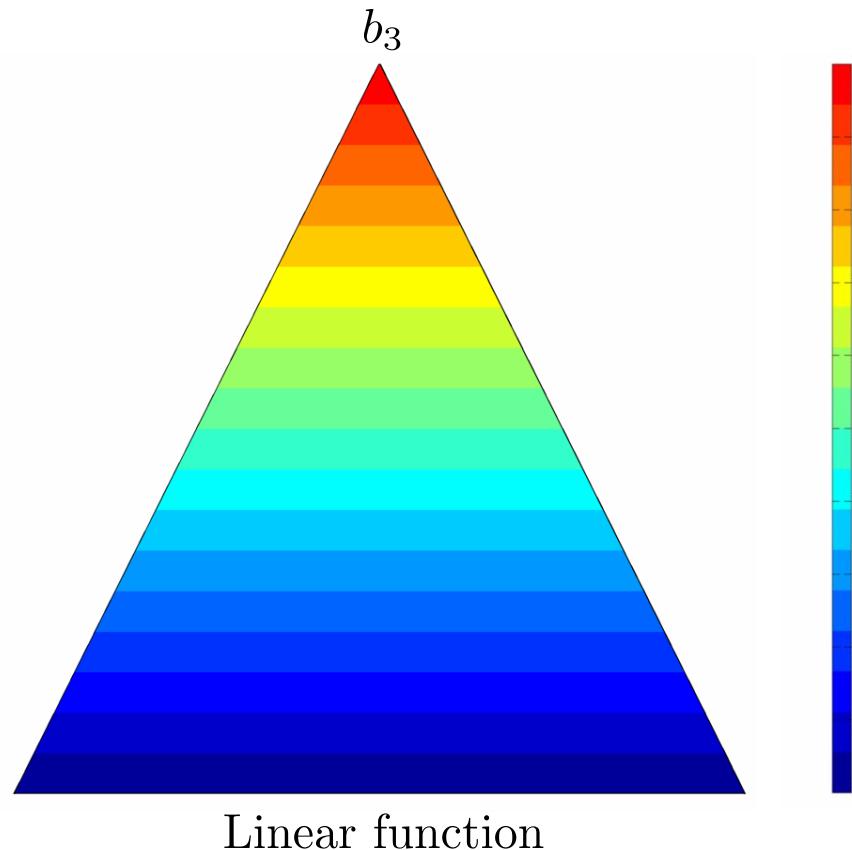
Closed-form:  $b_i$  are expressed in analytic form

# Triangle barycentric coordinates

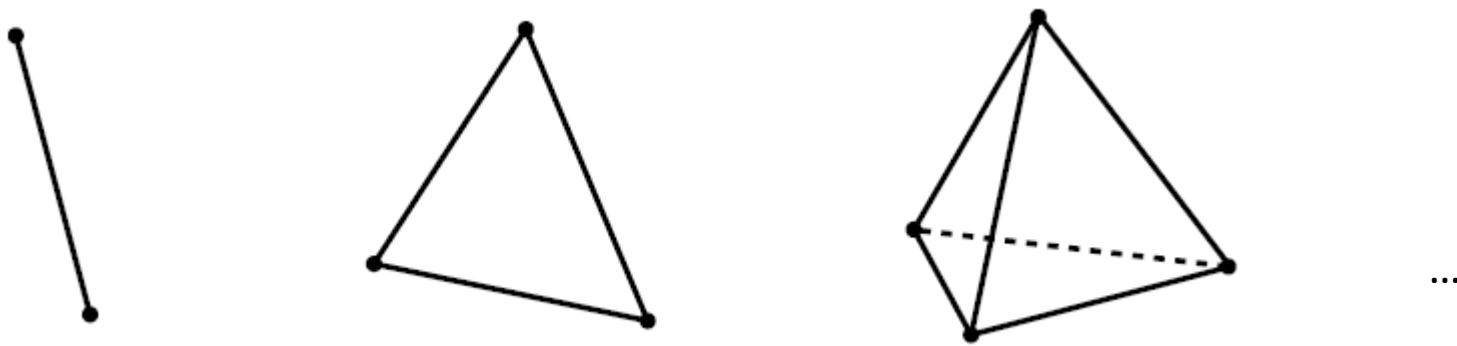
---

$$\sum_{i=1}^{d+1} b_i(p) v_i$$

Data interpolation



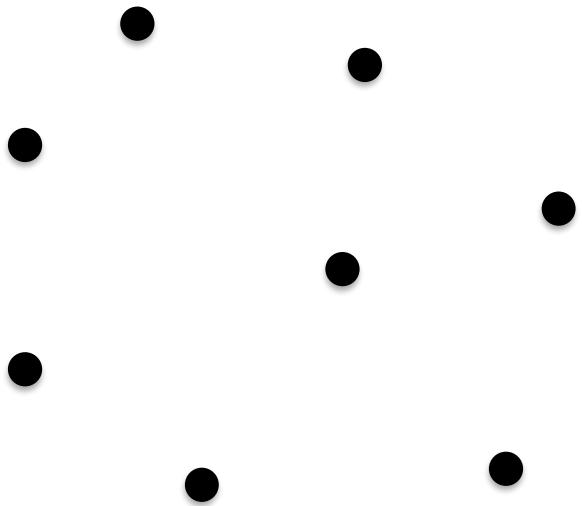
# Barycentric Coordinates for Simplices



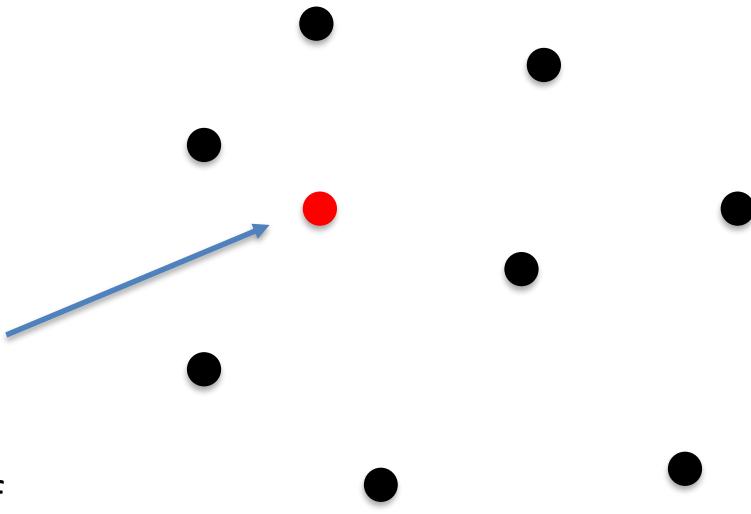
$$\sum_{i=1}^{d+1} b_i(p) v_i$$

# **BARYCENTRIC COORDINATES FOR POINT SETS**

# Point Set



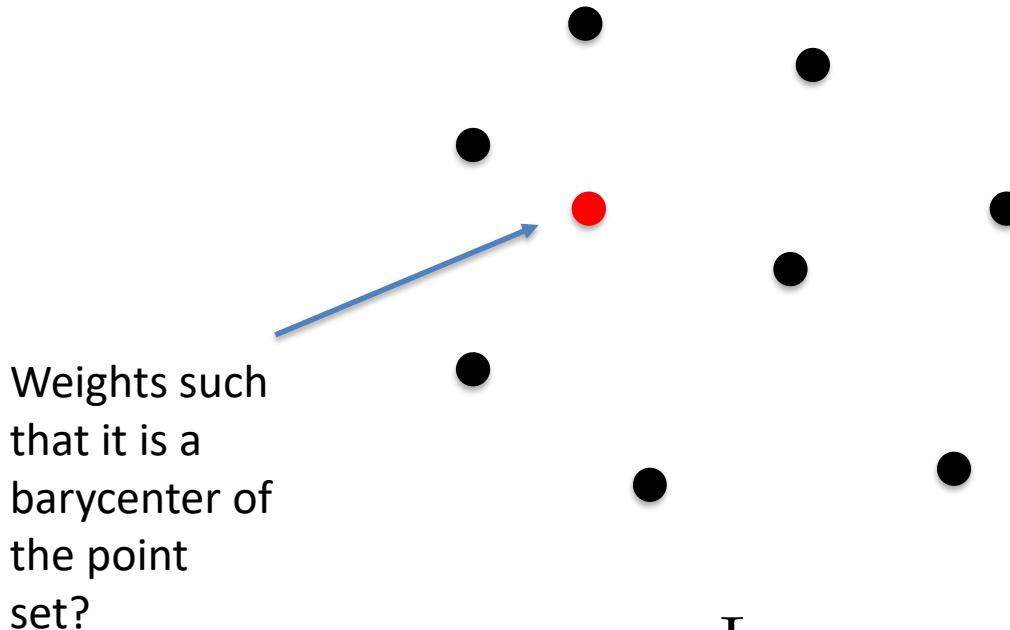
# Barycentric Coordinates?



Weights such  
that it is a  
barycenter of  
the point  
set?

Lagrange property:  $b_i(v_j) = \delta_{ij}$

# Barycentric Coordinates?



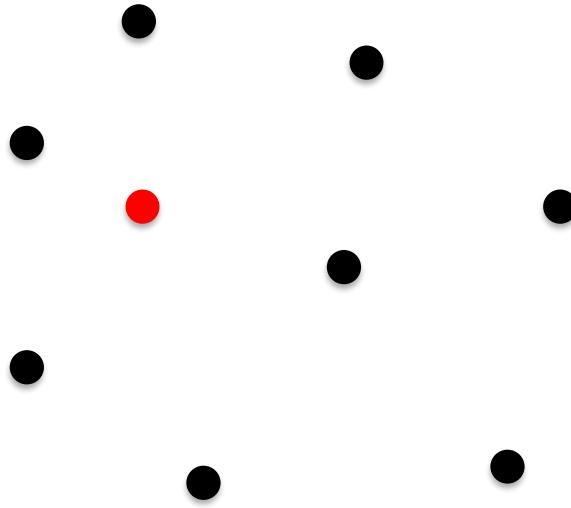
A. F. Möbius

[1790–1868]

Weights always exist if #points  $\geq$  dimension

$$\text{Lagrange property: } b_i(v_j) = \delta_{ij}$$

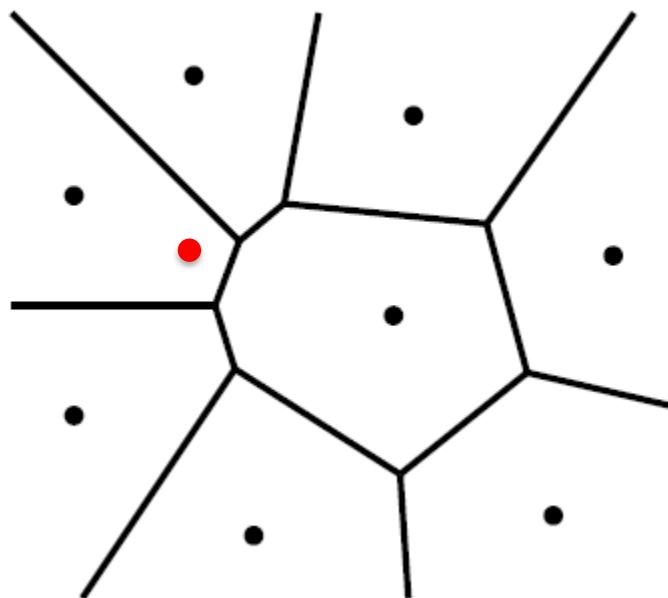
# Inverse Distance [Shepard]



$$u(\mathbf{x}) = \begin{cases} \frac{\sum_{i=1}^N w_i(\mathbf{x}) u_i}{\sum_{i=1}^N w_i(\mathbf{x})}, & \text{if } d(\mathbf{x}, \mathbf{x}_i) \neq 0 \text{ for all } i \\ u_i, & \text{if } d(\mathbf{x}, \mathbf{x}_i) = 0 \text{ for some } i \end{cases}$$

$$w_i(\mathbf{x}) = \frac{1}{d(\mathbf{x}, \mathbf{x}_i)^p}$$

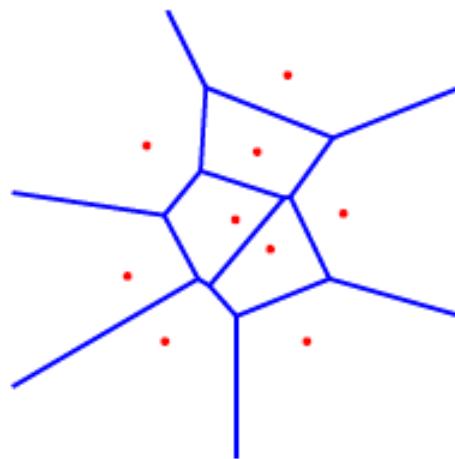
# Natural Neighbor Coordinates [Sibson]



# Voronoi Diagram

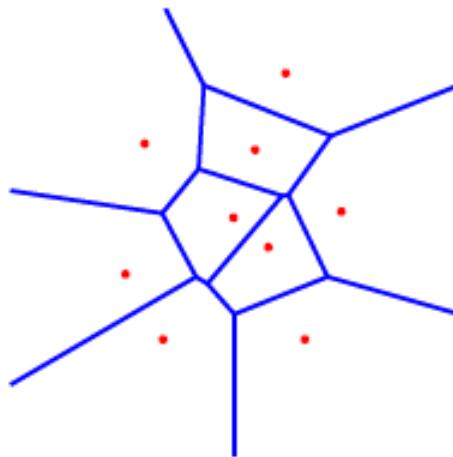
Let  $\mathcal{E} = \{\mathbf{p}_1, \dots, \mathbf{p}_n\}$  be a set of points (so-called sites) in  $\mathbb{R}^d$ . We associate to each site  $\mathbf{p}_i$  its Voronoi region  $V(\mathbf{p}_i)$  such that:

$$V(\mathbf{p}_i) = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x} - \mathbf{p}_i\| \leq \|\mathbf{x} - \mathbf{p}_j\|, \forall j \leq n\}.$$



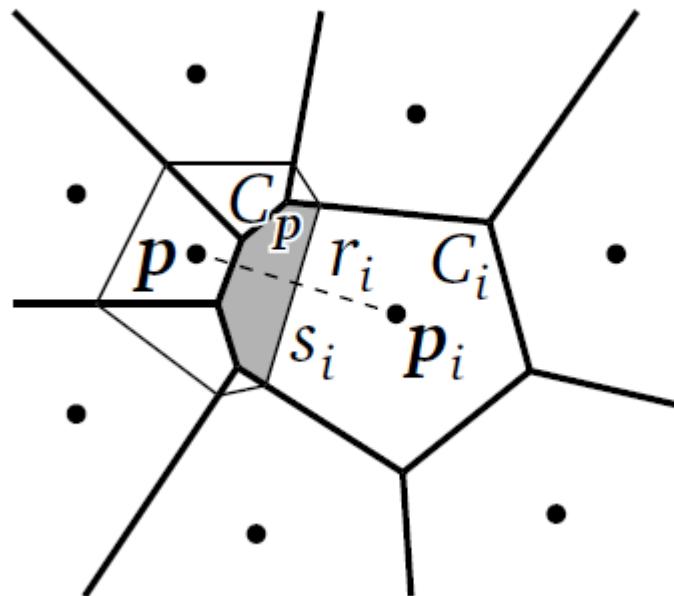
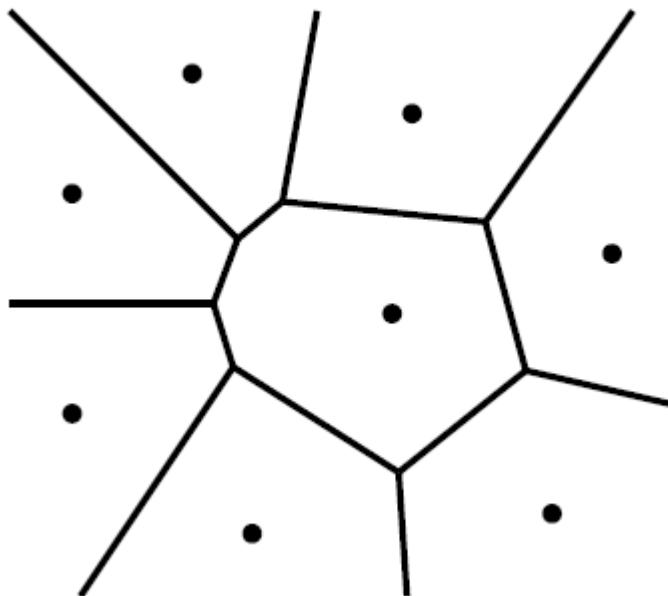
# Voronoi Diagram

- The collection of the non-empty Voronoi regions and their faces, together with their incidence relations, constitute a cell complex called the **Voronoi diagram** of E.
- The locus of points which are equidistant to two sites  $p_i$  and  $p_j$  is called a **bisector**, all bisectors being affine subspaces of  $\mathbb{R}^d$  (lines in 2D).



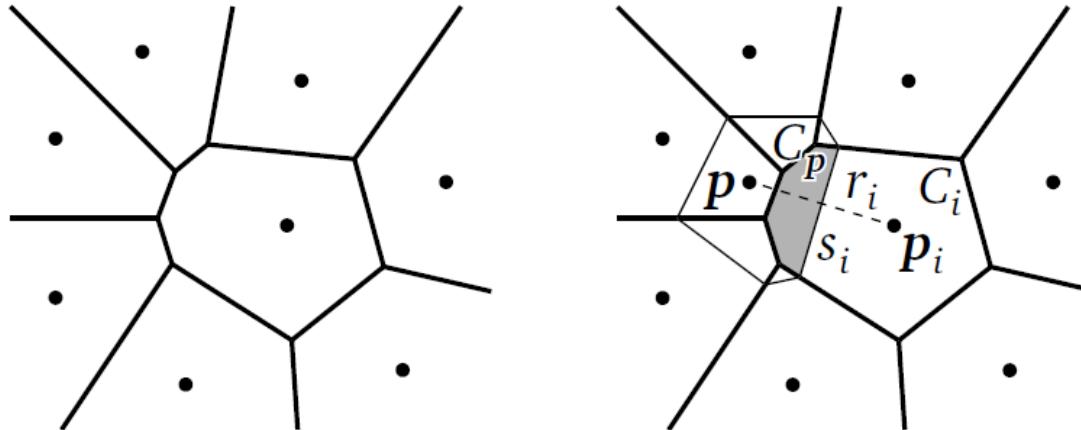
demo

# Sibson Coordinates



$$w_i = \text{Area}[C_i \cap C_p], \quad i = 1, \dots, n.$$

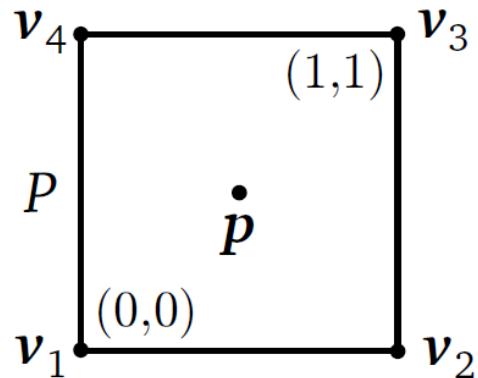
# Sibson Coordinates



- Well define over convex hull
- Local support
- Satisfy Lagrange property
- $C^1$  continuity, except at points  $p_i$  (only  $C^0$ )

# **GENERALIZED BARYCENTRIC COORDINATES**

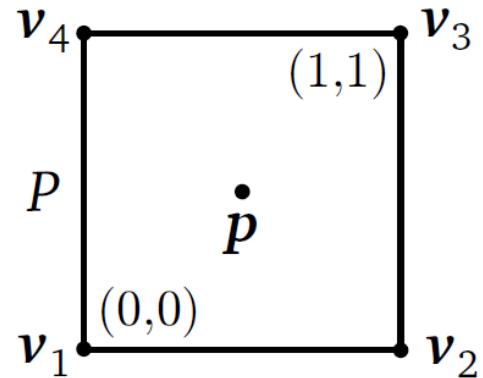
# Quadrilateral?



$$b_1(p) = (1-s)(1-t), \quad b_2(p) = s(1-t), \quad b_3(p) = st, \quad b_4(p) = (1-s)t$$

Bilinear interpolation

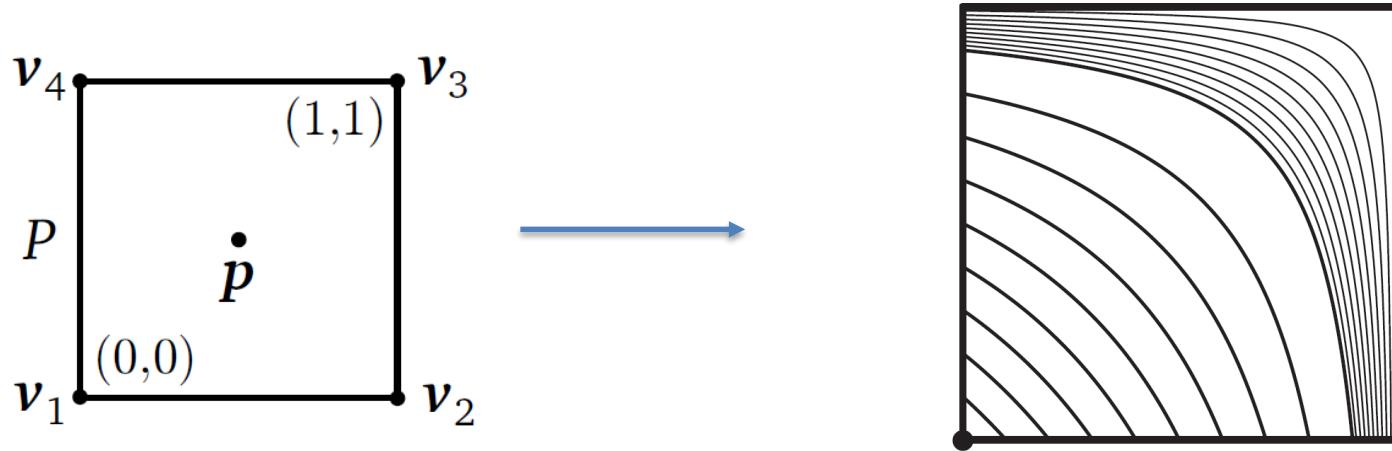
# Quadrilateral



$$b_1(p) = (1-s)(1-t), \quad b_2(p) = s(1-t), \quad b_3(p) = st, \quad b_4(p) = (1-s)t$$

Bilinear map on unit square

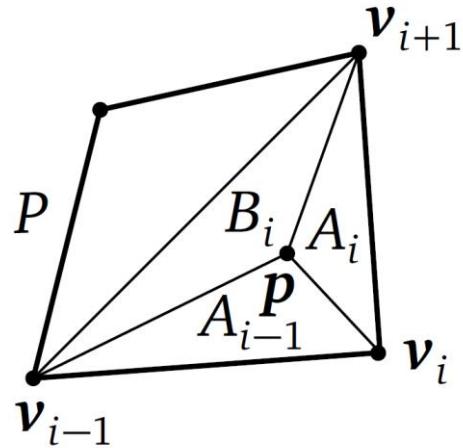
# Quadrilateral



$$b_1(p) = (1-s)(1-t), \quad b_2(p) = s(1-t), \quad b_3(p) = st, \quad b_4(p) = (1-s)t$$

Image of bilinear map on unit square

# Unified Formula! [Floater]

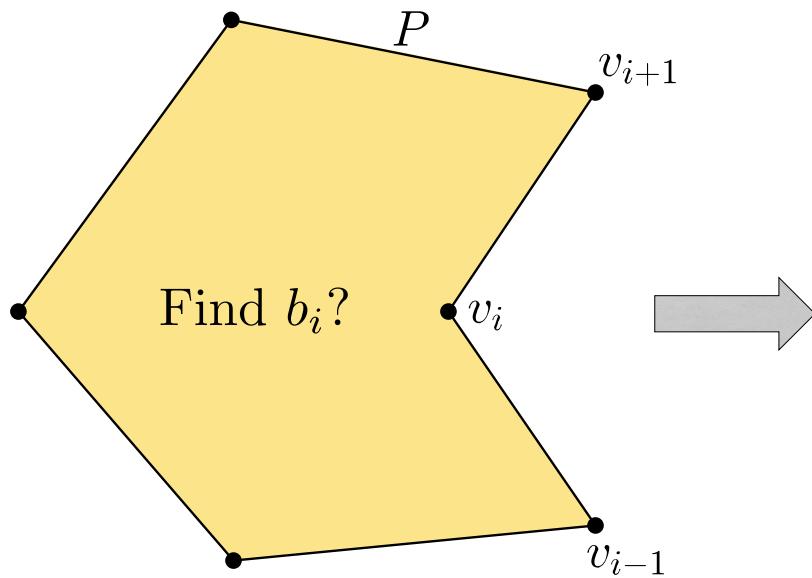


$$b_i(p) = \frac{4A_{i+1}(p)A_{i+2}(p)}{G_{i+1}(p)G_{i+2}(p)}, \quad i = 1, \dots, 4$$

$$G_i = 2A_i - B_i - B_{i+1} + \sqrt{B_1^2 + B_2^2 + 2A_1A_3 + 2A_2A_4},$$

signed area

# Generalized barycentric coordinates

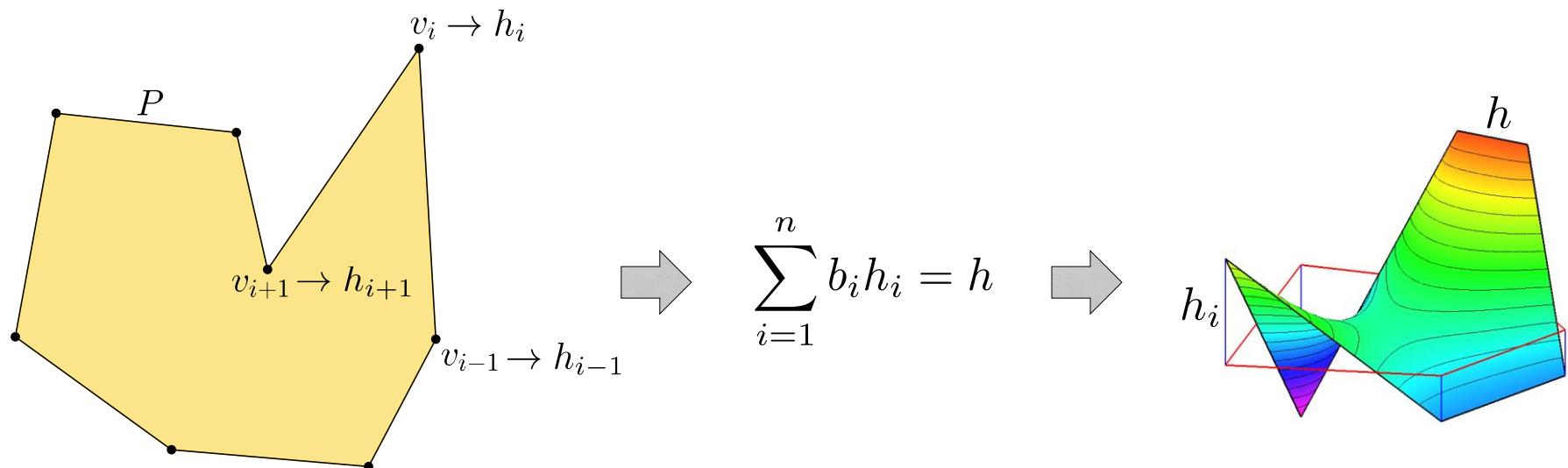


## Properties:

- Constant precision
- Linear precision
- The Lagrange property
- Non-negativity
- Linearity along edges
- Smoothness
- Closed-form

# Interpolation

---



# Generalized barycentric coordinates

---

$$b_i = \frac{w_i}{W}, \text{ where } W = \sum_{j=1}^n w_j$$

Different weights -> different coordinate functions

# Generalized barycentric coordinates

---

- Wachspress coordinates [Wachspress, 1975]
- Discrete harmonic coordinates [Pinkall and Polthier, 1993]
- Mean value coordinates [Floater, 2003]
- Metric coordinates [Malsch et al., 2005]
- Harmonic coordinates [Joshi et al., 2007]
- Maximum entropy coordinates [Hormann and Sukumar, 2008]
- Complex barycentric coordinates [Weber et al., 2009]
- Moving least squares coordinates [Manson and Schaefer, 2010]
- Cubic mean value coordinates [Li and Hu, 2013]
- Poisson coordinates [Li et al., 2013]

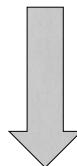
# Generalized barycentric coordinates

---

- Wachspress coordinates [Wachspress, 1975]
- Discrete harmonic coordinates [Pinkall and Polthier, 1993]
- Mean value coordinates [Floater, 2003]
- Metric coordinates [Malsch et al., 2005]
- Harmonic coordinates [Joshi et al., 2007]
- Maximum entropy coordinates [Hormann and Sukumar, 2008]
- Complex barycentric coordinates [Weber et al., 2009]
- Moving least squares coordinates [Manson and Schaefer, 2010]
- Cubic mean value coordinates [Li and Hu, 2013]
- Poisson coordinates [Li et al., 2013]

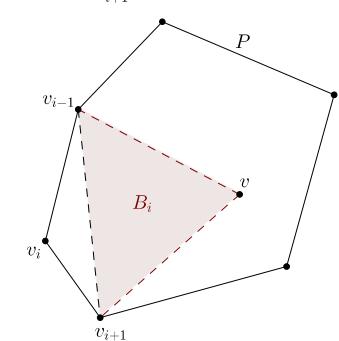
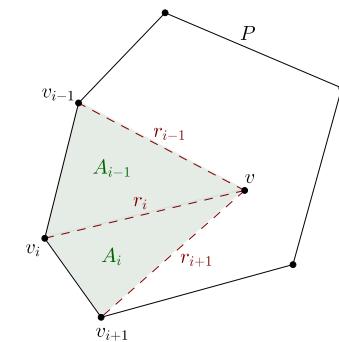
# Generalized barycentric coordinates

- Wachspress coordinates [Wachspress, 1975]
- Discrete harmonic coordinates [Pinkall and Polthier, 1993]
- Mean value coordinates [Floater, 2003]



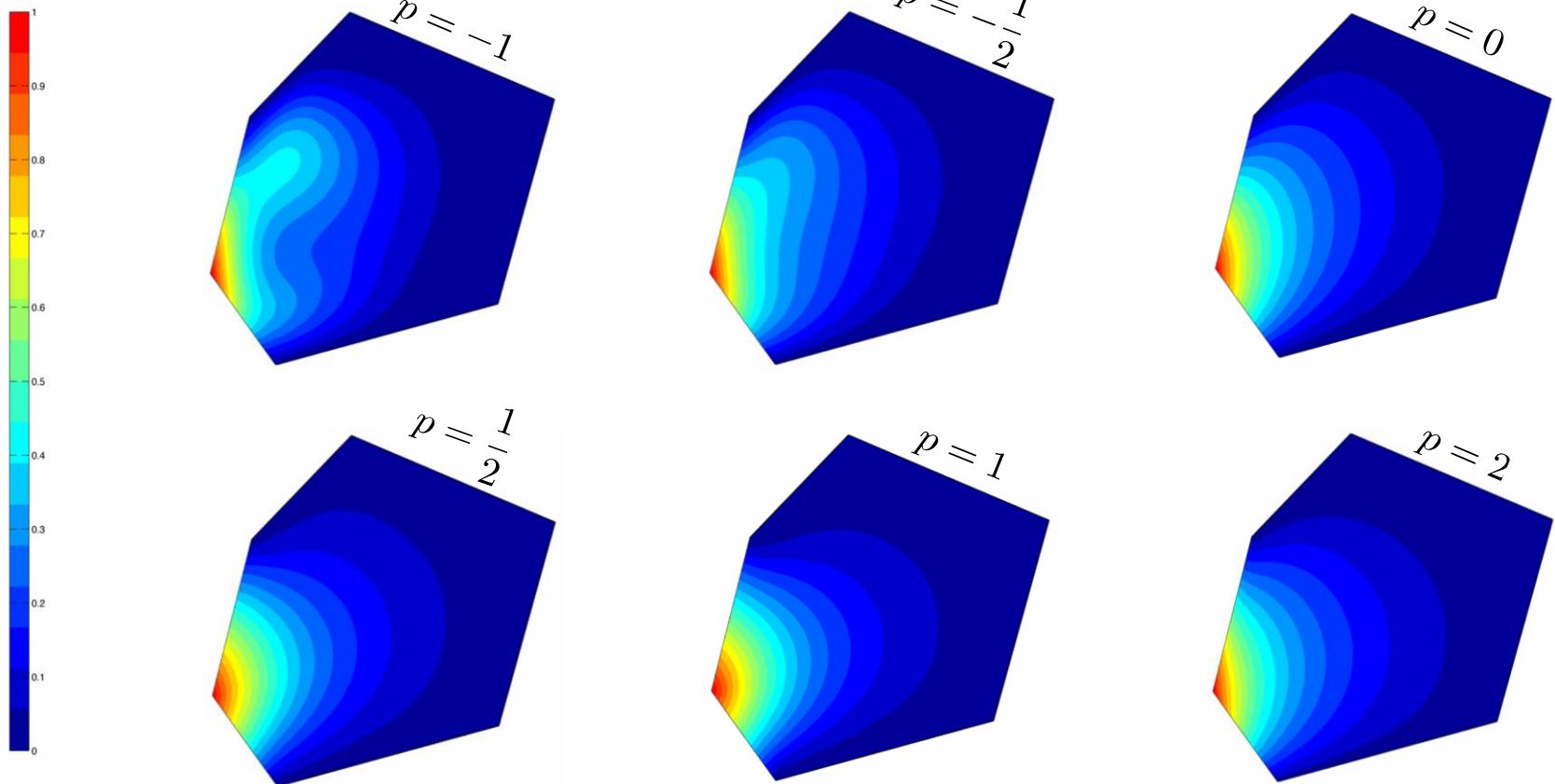
- Three-point coordinates [Floater et al., 2006]

$$p \in \mathbb{R}, \quad w_i = \frac{r_{i-1}^p A_i - r_i^p B_i + r_{i+1}^p A_{i-1}}{A_{i-1} A_i}$$



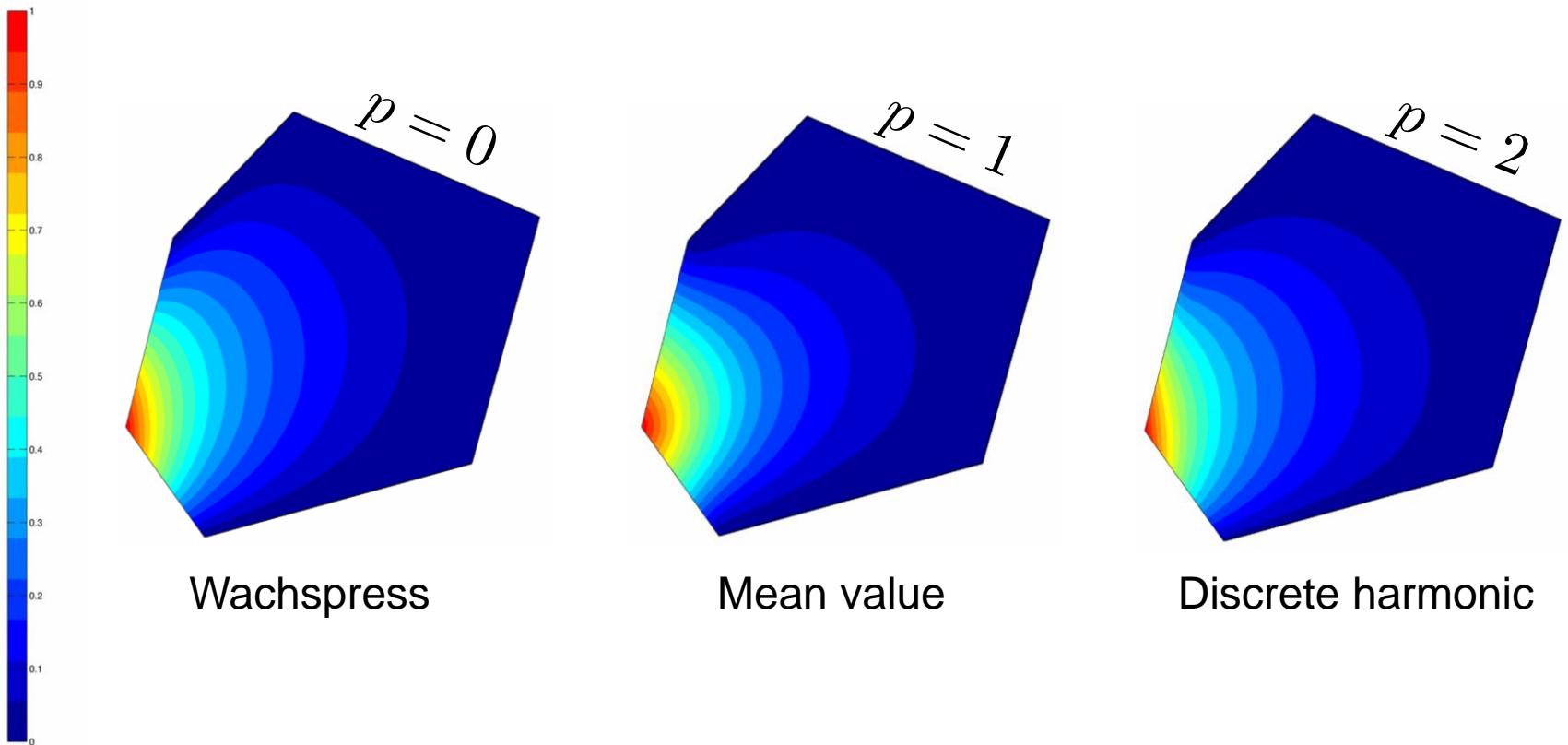
# Three-point coordinates

---



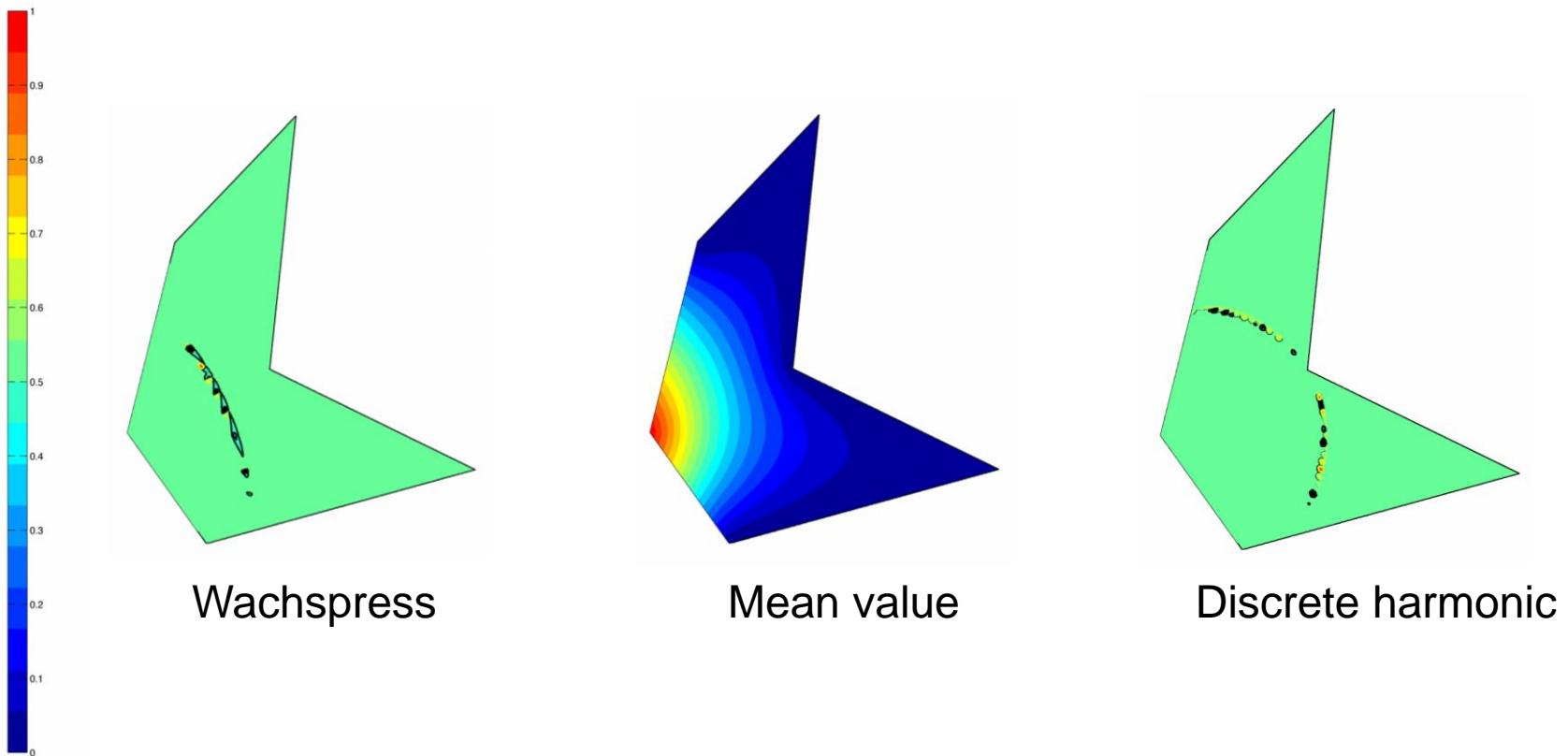
# Three-point coordinates

---



# Three-point coordinates

---



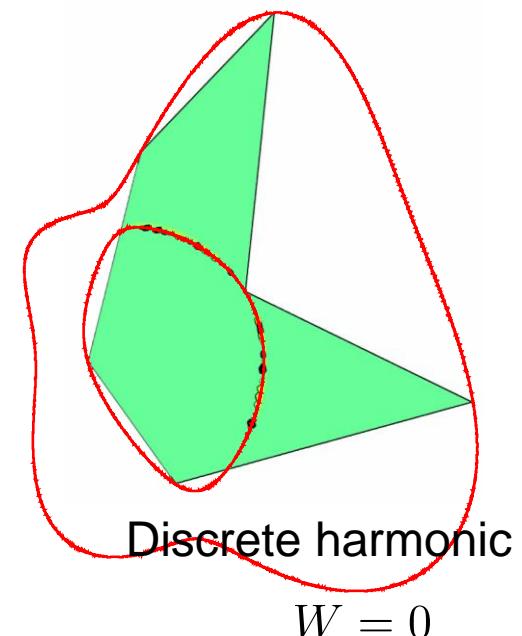
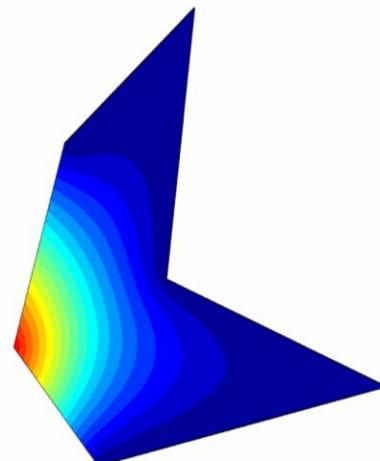
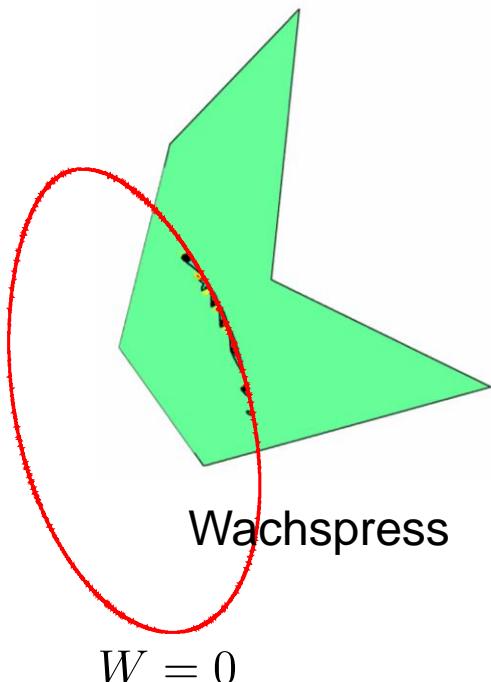
Wachspress

Mean value

Discrete harmonic

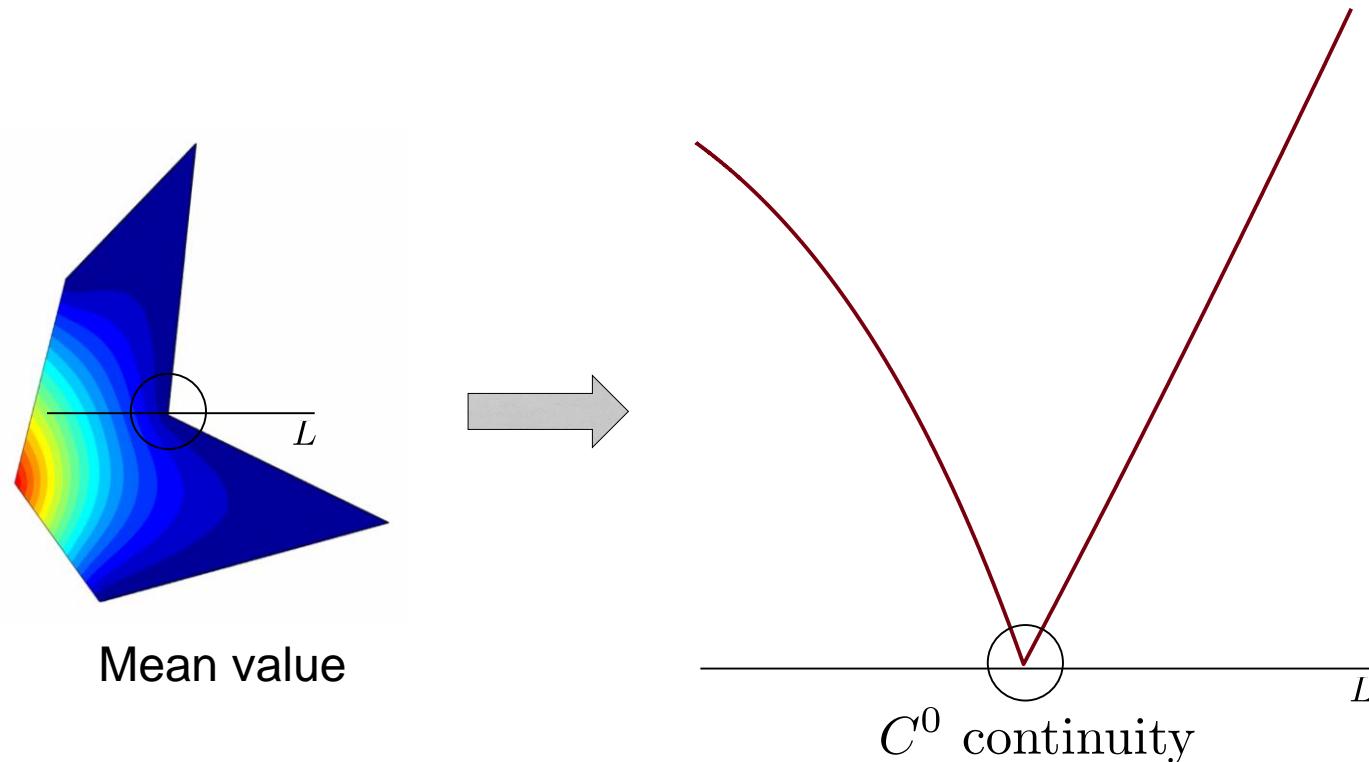
# Three-point coordinates

---



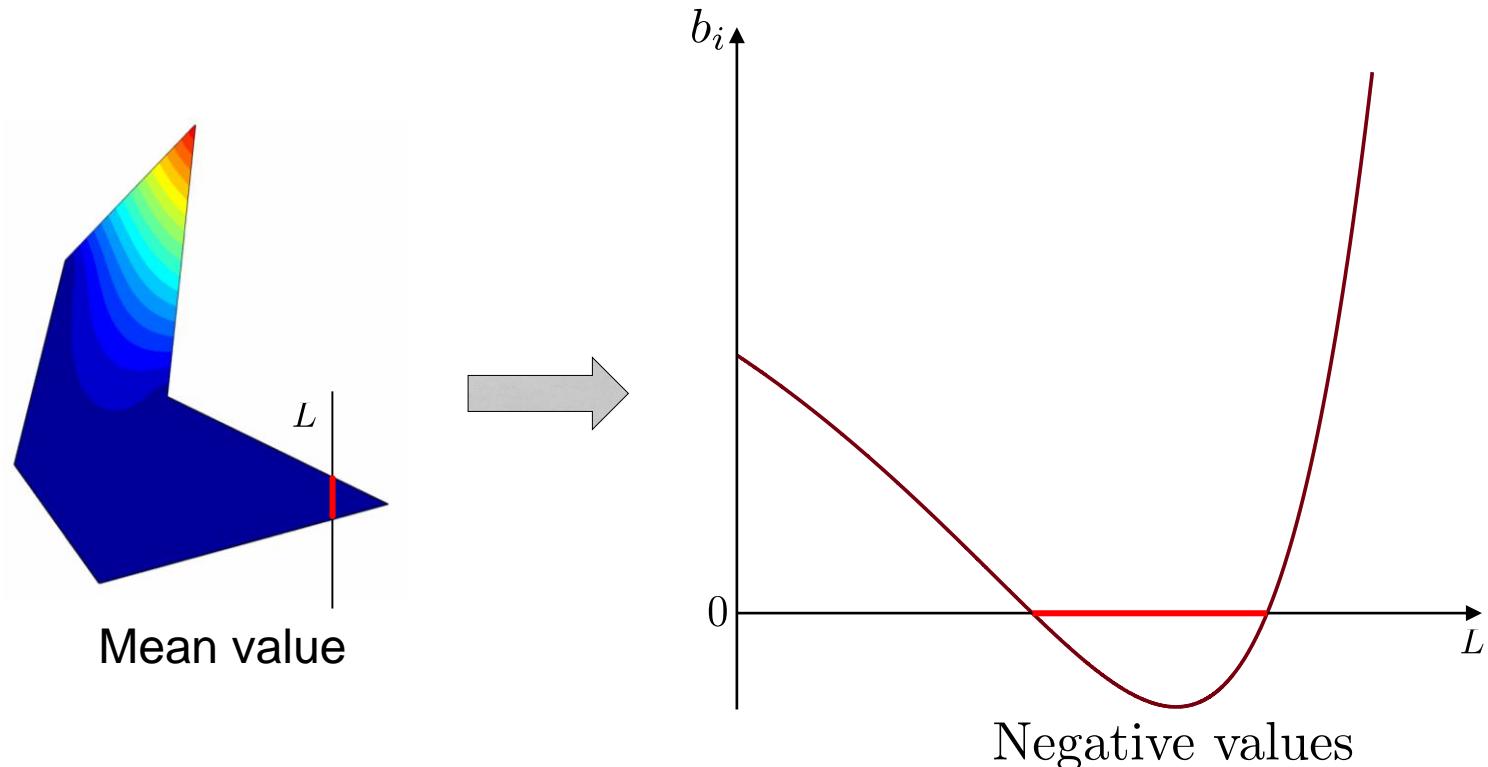
# Mean value coordinates

---



# Mean value coordinates

Positive mean value coordinates [Lipman et al., 2007]

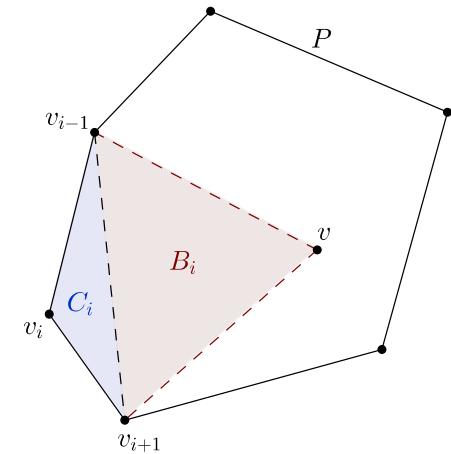
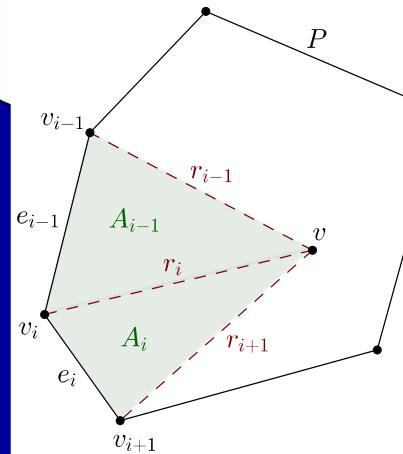
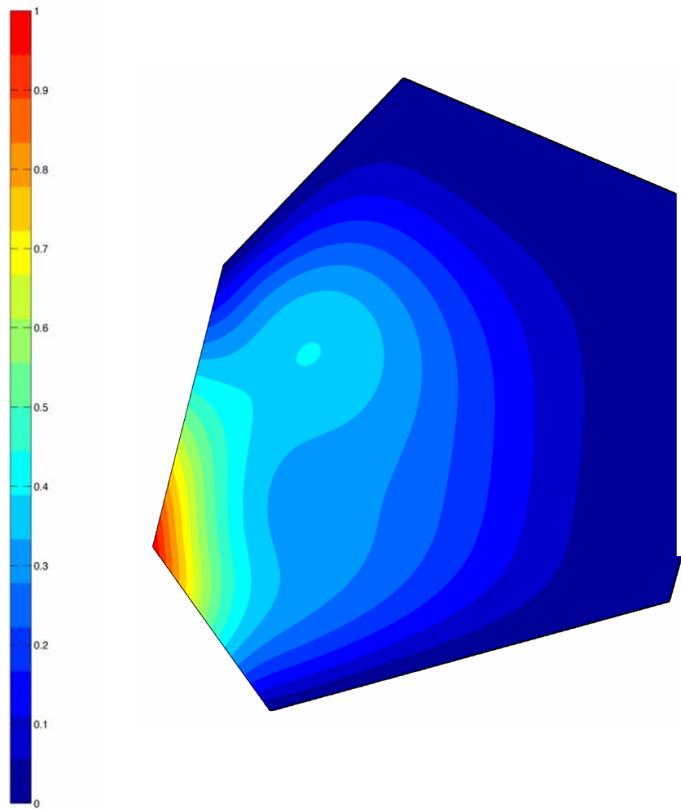


# Generalized barycentric coordinates

---

- Wachspress coordinates [Wachspress, 1975]
- Discrete harmonic coordinates [Pinkall and Polthier, 1993]
- Mean value coordinates [Floater, 2003]
- Metric coordinates [Malsch et al., 2005]
- Harmonic coordinates [Joshi et al., 2007]
- Maximum entropy coordinates [Hormann and Sukumar, 2008]
- Complex barycentric coordinates [Weber et al., 2009]
- Moving least squares coordinates [Manson and Schaefer, 2010]
- Cubic mean value coordinates [Li and Hu, 2013]
- Poisson coordinates [Li et al., 2013]

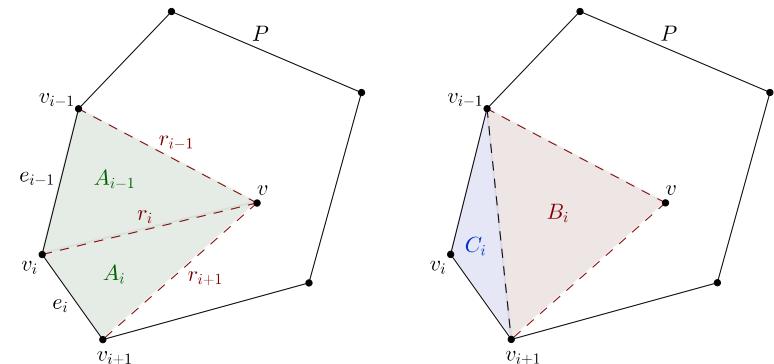
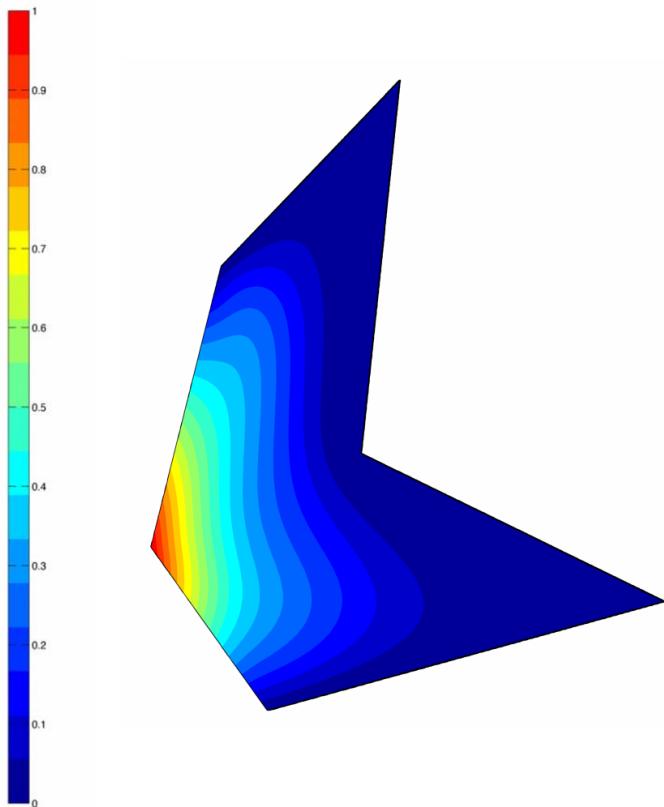
# Metric coordinates



$$q_i = r_i + r_{i+1} - e_i$$

$$w_i = \frac{A_{i-2}}{C_{i-1}q_{i-2}q_{i-1}} - \frac{B_i}{C_iq_{i-1}q_i} + \frac{A_{i+1}}{C_{i+1}q_iq_{i+1}}$$

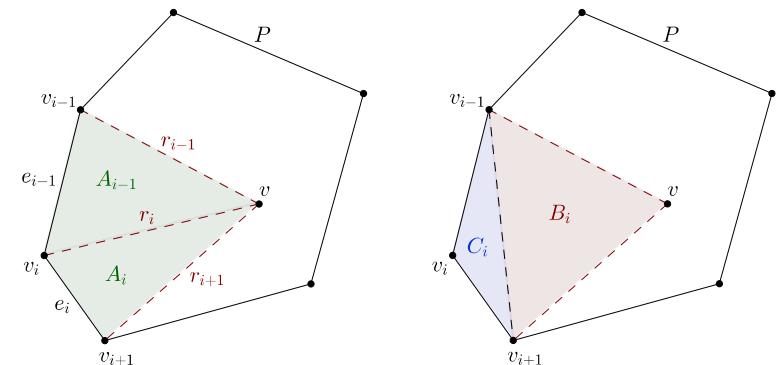
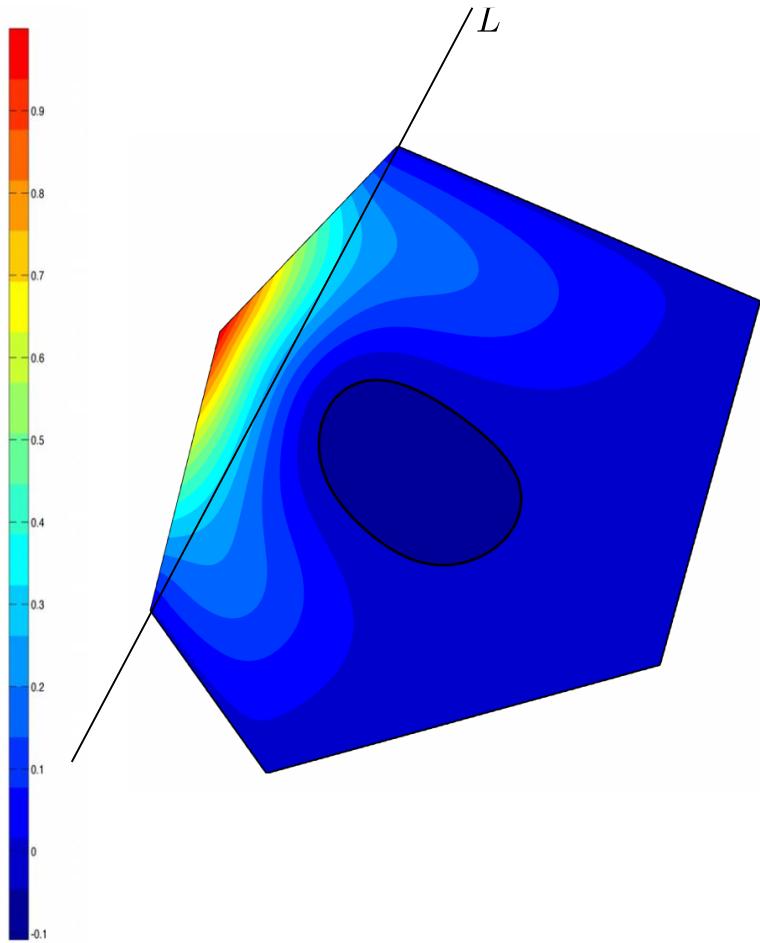
# Metric coordinates



$$q_i = r_i + r_{i+1} - e_i$$

$$w_i = \frac{A_{i-2}}{C_{i-1}q_{i-2}q_{i-1}} - \frac{B_i}{C_iq_{i-1}q_i} + \frac{A_{i+1}}{C_{i+1}q_iq_{i+1}}$$

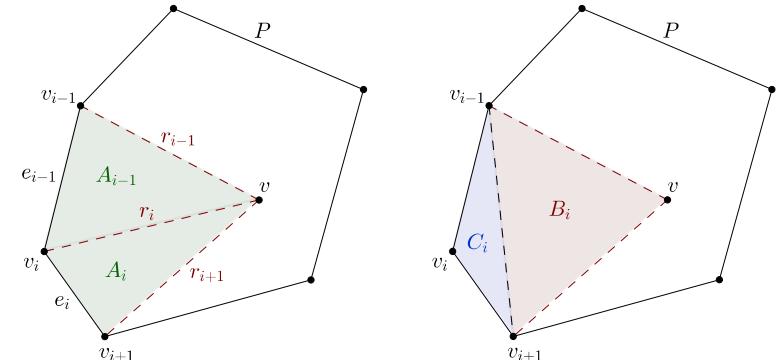
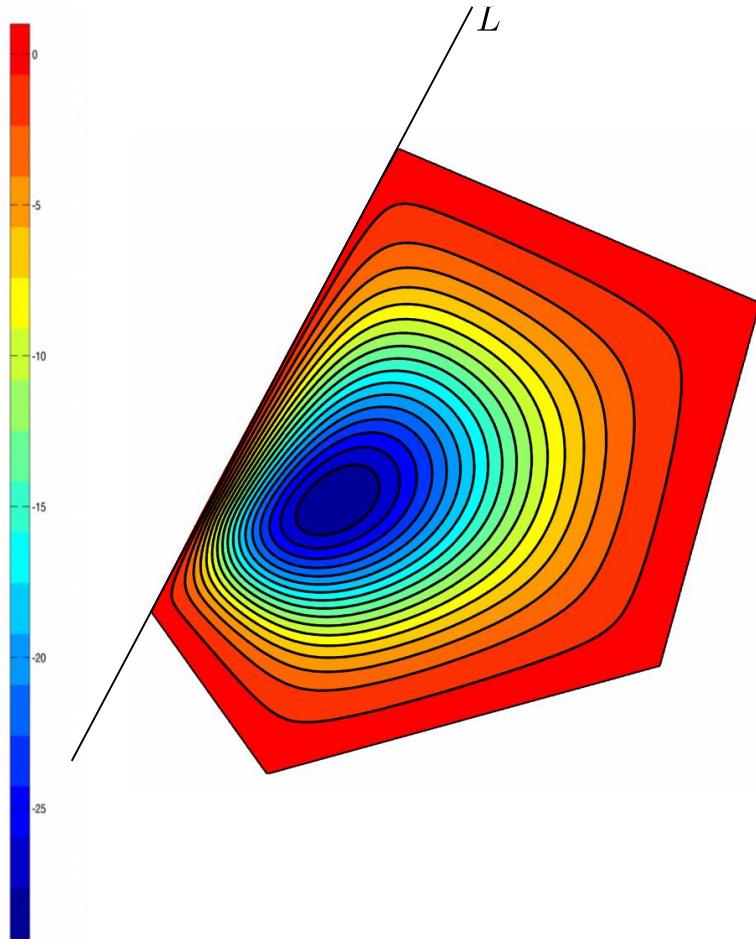
# Metric coordinates



$$q_i = r_i + r_{i+1} - e_i$$

$$w_i = \frac{A_{i-2}}{C_{i-1}q_{i-2}q_{i-1}} - \frac{B_i}{C_iq_{i-1}q_i} + \frac{A_{i+1}}{C_{i+1}q_iq_{i+1}}$$

# Metric coordinates



$$q_i = r_i + r_{i+1} - e_i$$

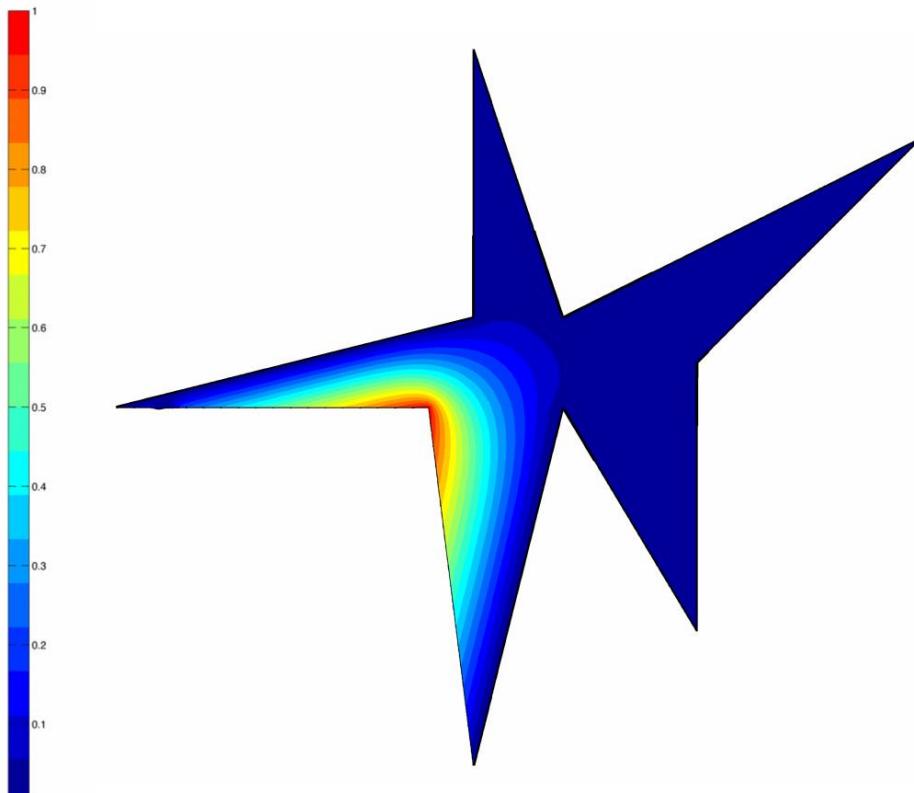
$$w_i = \frac{A_{i-2}}{C_{i-1}q_{i-2}q_{i-1}} - \frac{B_i}{C_iq_{i-1}q_i} + \frac{A_{i+1}}{C_{i+1}q_iq_{i+1}}$$

# Generalized barycentric coordinates

---

- Wachspress coordinates [Wachspress, 1975]
- Discrete harmonic coordinates [Pinkall and Polthier, 1993]
- Mean value coordinates [Floater, 2003]
- Metric coordinates [Malsch et al., 2005]
- **Harmonic coordinates** [Joshi et al., 2007]
- Maximum entropy coordinates [Hormann and Sukumar, 2008]
- Complex barycentric coordinates [Weber et al., 2009]
- Moving least squares coordinates [Manson and Schaefer, 2010]
- Cubic mean value coordinates [Li and Hu, 2013]
- Poisson coordinates [Li et al., 2013]

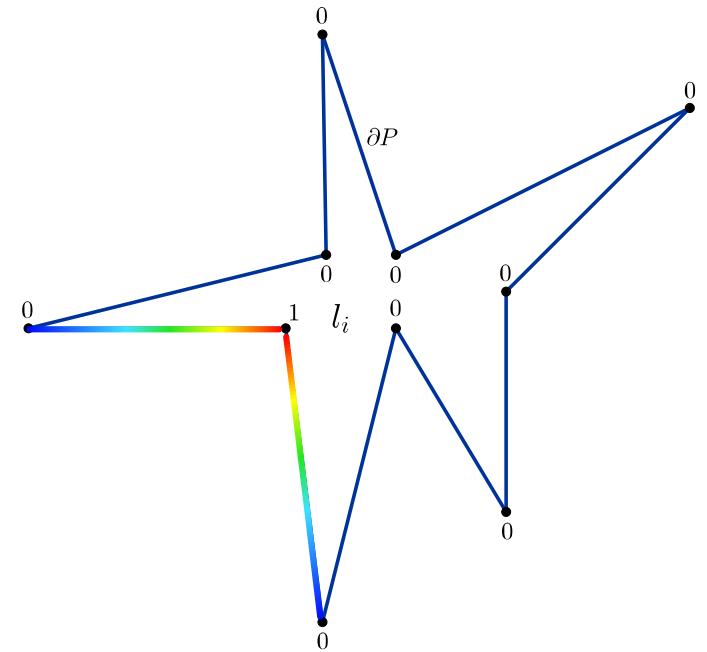
# Harmonic coordinates



No closed form!

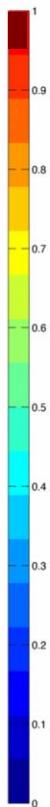
Solve the Laplacian equation:

$$\Delta b_i = 0 \text{ s.t. } b_i|_{\partial P} = l_i$$

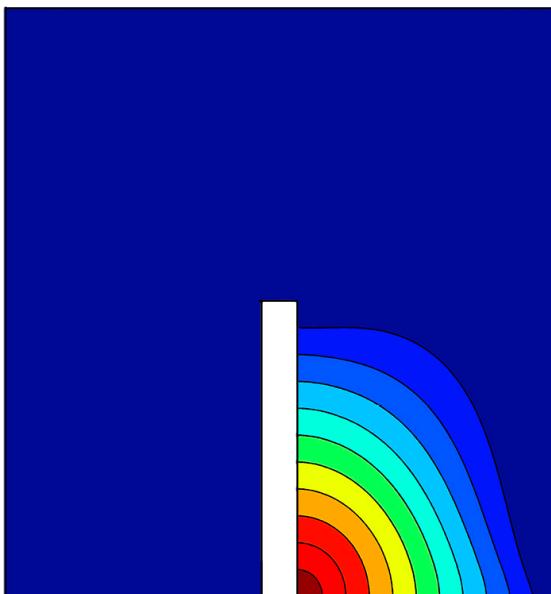


# Biharmonic vs Harmonic

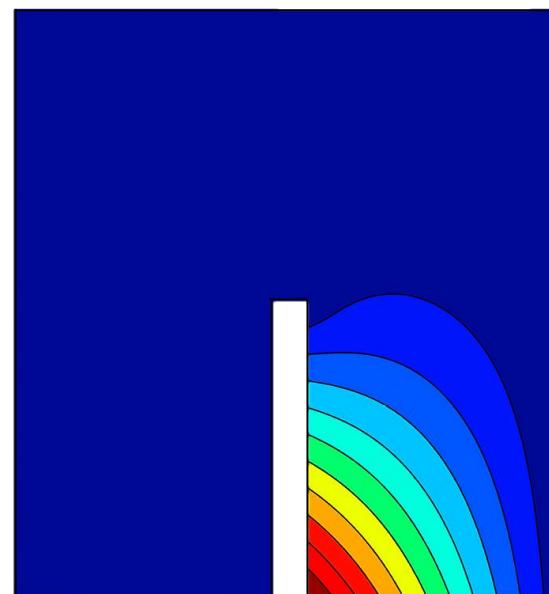
---



[Weber et al., 2012]



Biharmonic



Harmonic

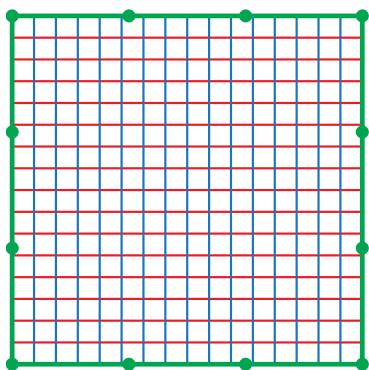
# Generalized barycentric coordinates

---

- Wachspress coordinates [Wachspress, 1975]
- Discrete harmonic coordinates [Pinkall and Polthier, 1993]
- Mean value coordinates [Floater, 2003]
- Metric coordinates [Malsch et al., 2005]
- Harmonic coordinates [Joshi et al., 2007]
- Maximum entropy coordinates [Hormann and Sukumar, 2008]
- Complex barycentric coordinates [Weber et al., 2009]
- Moving least squares coordinates [Manson and Schaefer, 2010]
- Cubic mean value coordinates [Li and Hu, 2013]
- Poisson coordinates [Li et al., 2013]

# Barycentric mapping

Source polygon



$$P = [v_1, \dots, v_n]$$

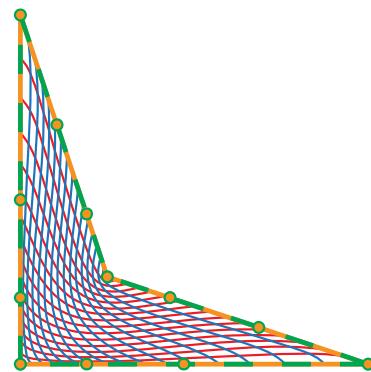
$$v_i \in \mathbb{R}^2$$

For  $v \in P$  :

$$f(v) = \sum_{i=1}^n b_i(v) \hat{v}_i$$



Target polygon

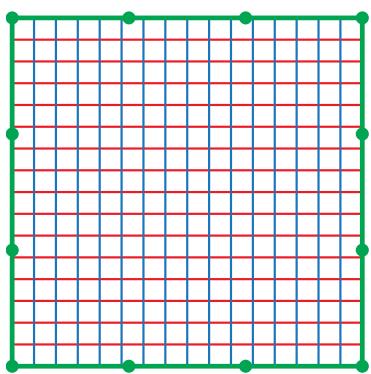


$$\hat{P} = [\hat{v}_1, \dots, \hat{v}_n]$$

$$\hat{v}_i \in \mathbb{R}^2$$

# Complex barycentric mapping

Source polygon



$$P = [z_1, \dots, z_n]$$

$$z_i \in \mathbb{C}$$

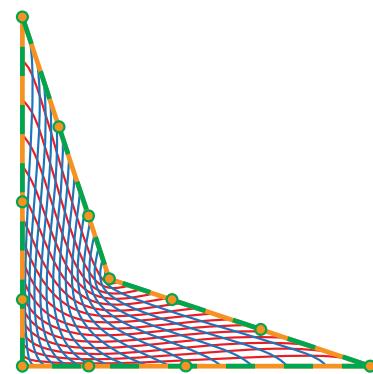
For  $z \in P$  :

$$g(z) = \sum_{i=1}^n c_i(z) \hat{z}_i$$



with complex  
barycentric coordinates  
 $c_i : P \rightarrow \mathbb{C}$

Target polygon



$$\hat{P} = [\hat{z}_1, \dots, \hat{z}_n]$$

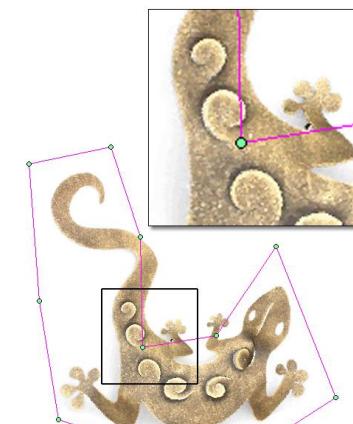
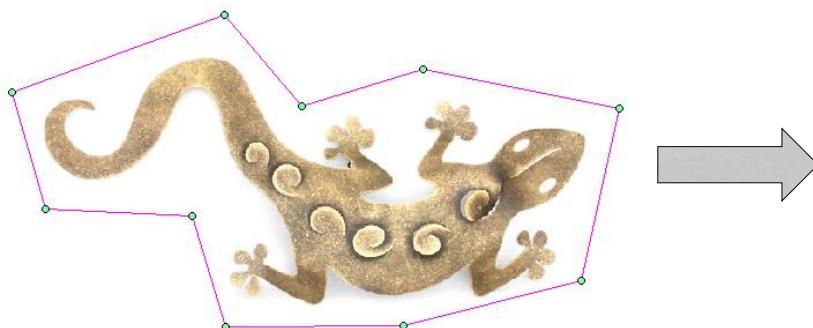
$$\hat{z}_i \in \mathbb{C}$$

# Complex barycentric coordinates

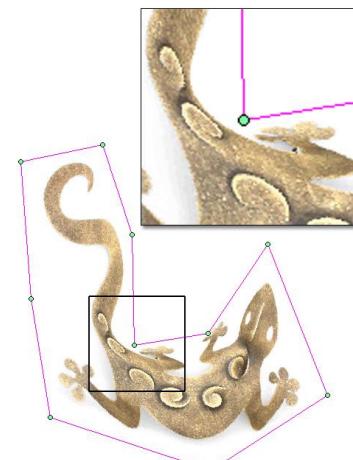
Three-point coordinates are generalized to complex three-point coordinates

Green coordinates are members of complex three-point coordinates  
[Lipman et al., 2008]

Induce conformal mappings



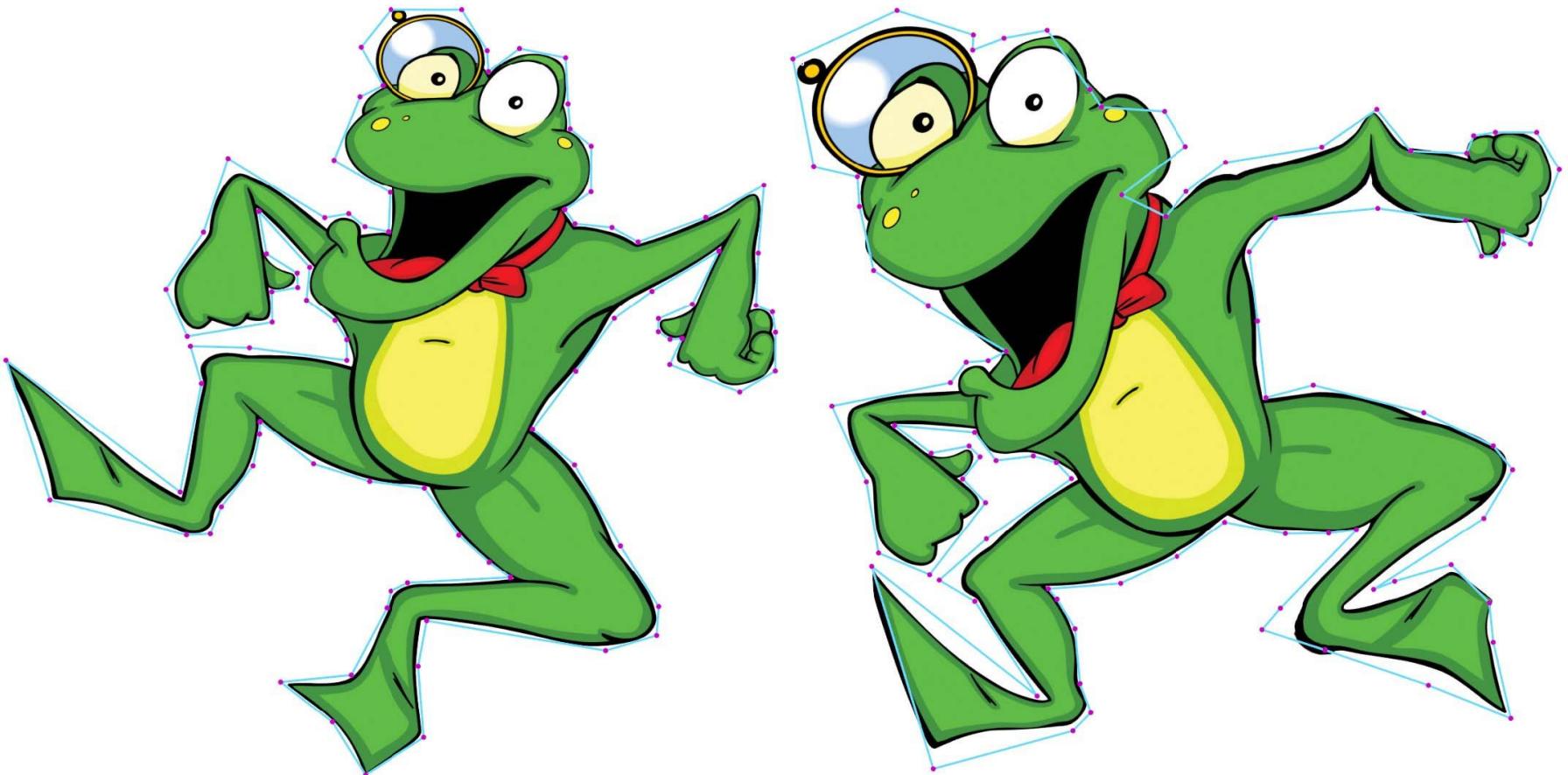
Green



Harmonic

# **APPLICATIONS**

# Editing

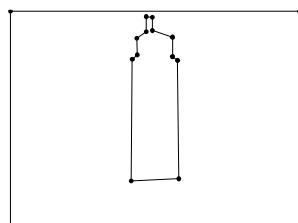
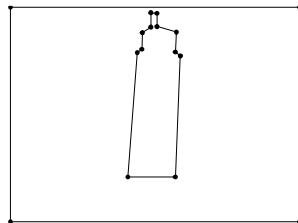


# Image warping

---



Original image



Mask



Warped image



# Principal Component Analysis

Linear Least Squares Approximation

Pierre Alliez  
Inria Sophia Antipolis

## Definition (point set case)

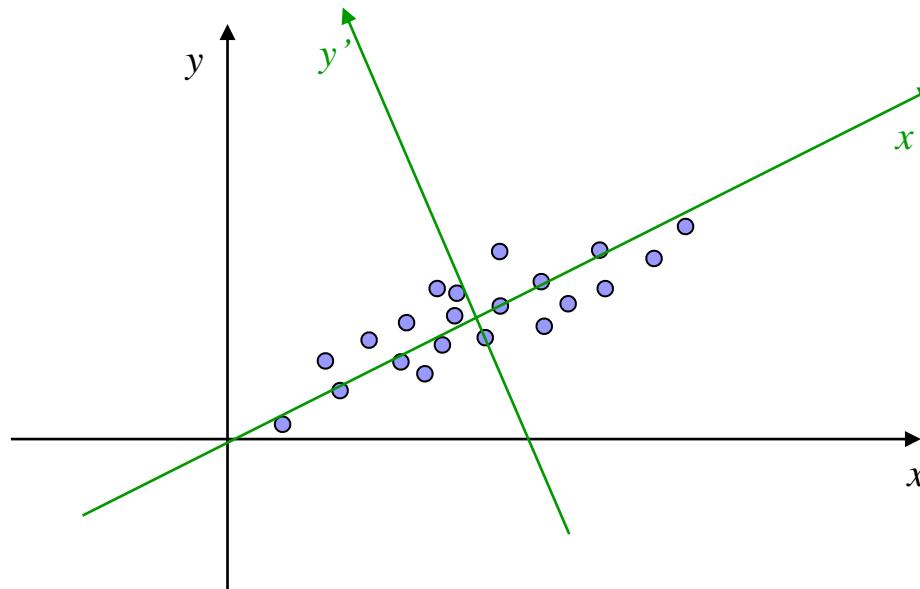
- Given a point set  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in R^d$ , linear least squares fitting amounts to find the linear sub-space of  $R^d$  which **minimizes the sum of squared distances** from the points to their projection onto this linear sub-space.

## Definition (point set case)

- This problem is equivalent to search for the linear sub-space which **maximizes the variance of projected points**, the latter being obtained by eigen decomposition of the covariance (scatter) matrix.
- Eigenvectors corresponding to large eigenvalues are the directions in which the data has **strong component**, or equivalently large variance. If eigenvalues are the same there is no preferable sub-space.

# PCA – the general idea

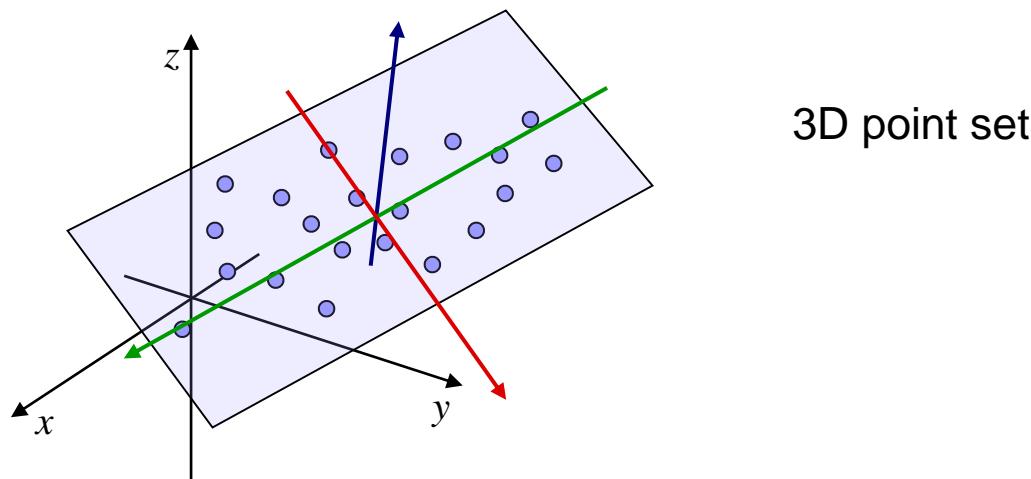
- PCA finds an orthogonal basis that best represents given data set.



- The sum of distances<sup>2</sup> from the  $x'$  axis is minimized.

# PCA – the general idea

- PCA finds an orthogonal basis that best represents given data set.



- PCA finds a best approximating plane (in terms of  $\sum distances^2$ )

# Notations

- Denote our data points by  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in R^d$

$$\mathbf{x}_1 = \begin{pmatrix} x_1^1 \\ x_1^2 \\ \vdots \\ x_1^d \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} x_2^1 \\ x_2^2 \\ \vdots \\ x_2^d \end{pmatrix}, \quad \dots, \quad \mathbf{x}_n = \begin{pmatrix} x_n^1 \\ x_n^2 \\ \vdots \\ x_n^d \end{pmatrix}$$

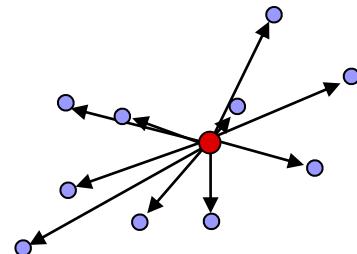
# The origin of the new axes

- The origin is zero-order approximation of our data set (a point)
- It will be the center of mass:

$$\mathbf{m} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$$

- It can be shown that:

$$\mathbf{m} = \underset{\mathbf{x}}{\operatorname{argmin}} \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{x}\|^2$$



# Scatter matrix

- Denote  $\mathbf{y}_i = \mathbf{x}_i - \mathbf{m}$ ,  $i = 1, 2, \dots, n$

$$S = YY^T$$

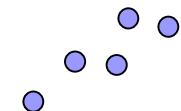
where  $\mathbf{Y}$  is  $d \times n$  matrix with  $\mathbf{y}_k$  as columns ( $k = 1, 2, \dots, n$ )

$$S = \begin{pmatrix} y_1^1 & y_2^1 & \vdots & y_n^1 \\ y_1^2 & y_2^2 & & y_n^2 \\ \vdots & \vdots & & \vdots \\ y_1^d & y_2^d & \vdots & y_n^d \end{pmatrix} \begin{pmatrix} y_1^1 & y_1^2 & \vdots & y_1^d \\ y_2^1 & y_2^2 & \vdots & y_2^d \\ \vdots & \vdots & & \vdots \\ y_n^1 & y_n^2 & \vdots & y_n^d \end{pmatrix}$$
$$\mathbf{Y} \quad \mathbf{Y}^T$$

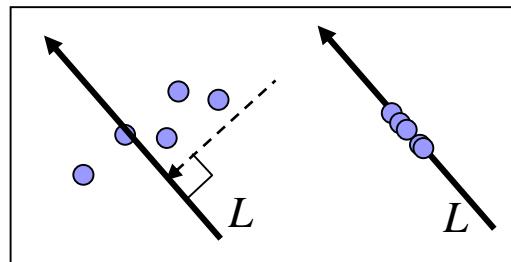
# Variance of projected points

- In a way,  $S$  measures variance (= scatterness) of the data in different directions.
- Let's look at a line  $L$  through the center of mass  $\mathbf{m}$ , and project our points  $\mathbf{x}_i$  onto it. The **variance of the projected points**  $\mathbf{x}'_i$  is:

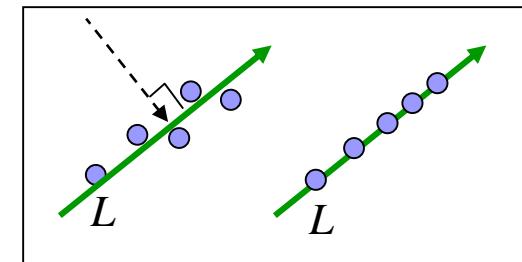
$$\text{var}(L) = \frac{1}{n} \sum_{i=1}^n \| \mathbf{x}'_i - \mathbf{m} \|^2$$



Original set



Small variance

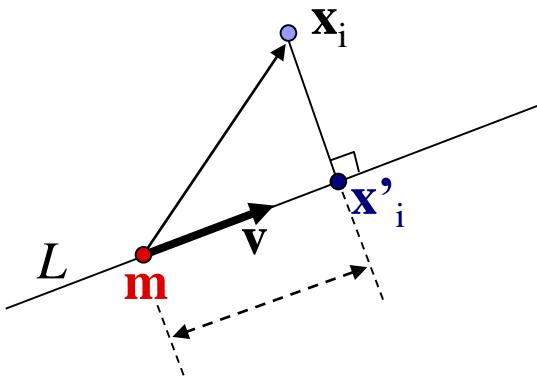


Large variance

# Variance of projected points

- Given a direction  $\mathbf{v}$ ,  $\|\mathbf{v}\| = 1$ , the projection of  $\mathbf{x}_i$  onto  $L = \mathbf{m} + \mathbf{v}t$  is:

$$\|\mathbf{x}'_i - \mathbf{m}\| = \langle \mathbf{v}, \mathbf{x}_i - \mathbf{m} \rangle \|\mathbf{v}\| = \langle \mathbf{v}, \mathbf{y}_i \rangle = \mathbf{v}^T \mathbf{y}_i$$



# Variance of projected points

■ So,

$$\begin{aligned}\text{var}(L) &= \frac{1}{n} \sum_{i=1}^n \| \mathbf{x}_i' \cdot \mathbf{m} \|^2 = \frac{1}{n} \sum_{i=1}^n (\mathbf{v}^T \mathbf{y}_i)^2 = \frac{1}{n} \| \mathbf{v}^T Y \|^2 = \\ &= \frac{1}{n} \| Y^T \mathbf{v} \|^2 = \frac{1}{n} \langle Y^T \mathbf{v}, Y^T \mathbf{v} \rangle = \frac{1}{n} \mathbf{v}^T Y Y^T \mathbf{v} = \frac{1}{n} \mathbf{v}^T S \mathbf{v} = \frac{1}{n} \langle S \mathbf{v}, \mathbf{v} \rangle\end{aligned}$$

$$\sum_{i=1}^n (\mathbf{v}^T \mathbf{y}_i)^2 = \sum_{i=1}^n \left( \begin{pmatrix} v^1 & v^2 & \dots & v^d \end{pmatrix} \begin{pmatrix} y_i^1 \\ y_i^2 \\ \vdots \\ y_i^d \end{pmatrix} \right)^2 = \left\| \begin{pmatrix} v^1 & v^2 & \dots & v^d \end{pmatrix} \begin{pmatrix} y_1^1 & y_2^1 & \dots & y_n^1 \\ y_1^2 & y_2^2 & \dots & y_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ y_1^d & y_2^d & \dots & y_n^d \end{pmatrix} \right\|^2 = \| \mathbf{v}^T Y \|^2$$

# Directions of maximal variance

- So, we have:  $\text{var}(L) = \langle S\mathbf{v}, \mathbf{v} \rangle$
- Theorem:

Let  $f: \{\mathbf{v} \in R^d \mid \|\mathbf{v}\| = 1\} \rightarrow R$ ,

$f(\mathbf{v}) = \langle S\mathbf{v}, \mathbf{v} \rangle$  (and  $S$  is a symmetric matrix).

Then, the extrema of  $f$  are attained at the eigenvectors of  $S$ .

- So, eigenvectors of  $S$  are directions of maximal/minimal variance.

# Summary so far

- We take the centered data points  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n \in R^d$
- Construct the scatter matrix  $S = YY^T$
- $S$  measures the variance of the data points
- Eigenvectors of  $S$  are directions of max/min variance.

# Scatter matrix - eigendecomposition

- $S$  is symmetric

$\Rightarrow S$  has eigendecomposition:  $S = V\Lambda V^T$

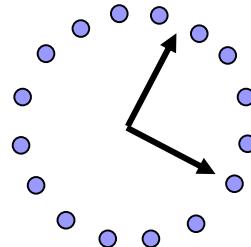
$$S = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_d \end{pmatrix}^T$$

The eigenvectors form  
orthogonal basis

# Principal components

- Eigenvectors that correspond to **big** eigenvalues are the directions in which the data has strong components (= large variance).
- If the eigenvalues are more or less the same – there is no preferable direction.

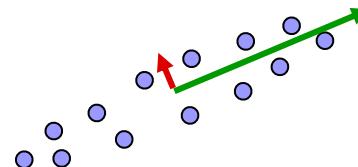
# Principal components



- There's no preferable direction
- $S$  looks like this:

$$\begin{pmatrix} \lambda & \\ & \lambda \end{pmatrix}$$

- Any vector is an eigenvector

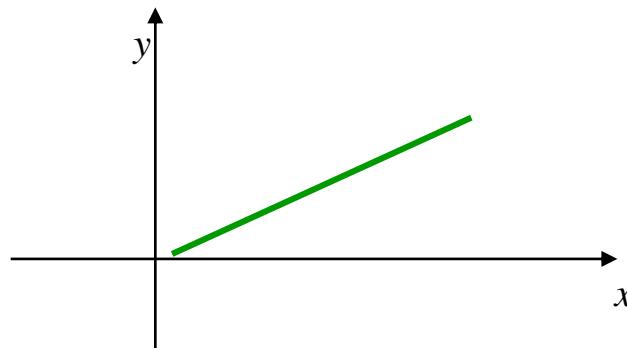
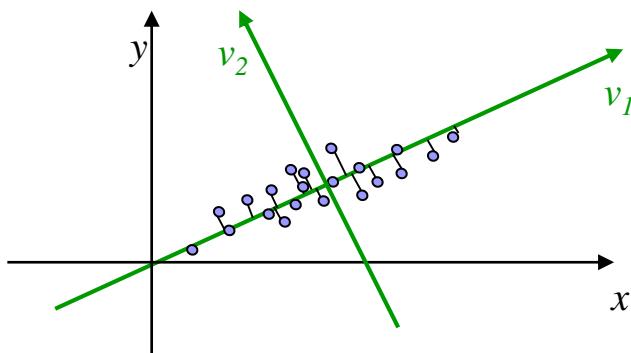


- There is a clear preferable direction
- $S$  looks like this:

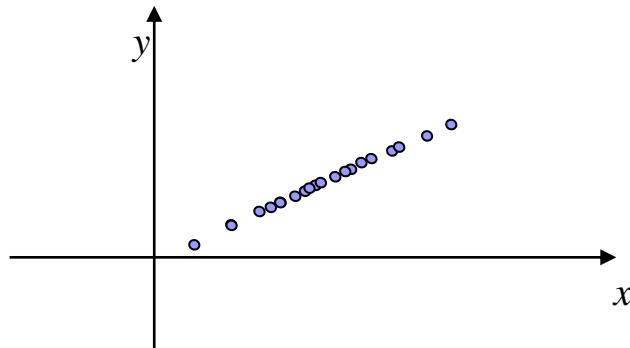
$$V \begin{pmatrix} \lambda & \\ & \mu \end{pmatrix} V^T$$

- $\mu$  is close to zero, much smaller than  $\lambda$ .

# For approximation



This line segment approximates  
the original data set



The projected data set  
approximates the original data  
set

# For approximation

- In general dimension  $d$ , the eigenvalues are sorted in descending order:

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$$

- The eigenvectors are sorted accordingly.
- To get an approximation of dimension  $d' < d$ , we take the  $d'$  first eigenvectors and look at the subspace they span ( $d' = 1$  is a line,  $d' = 2$  is a plane...)

# For approximation

- To get an approximating set, we project the original data points onto the chosen subspace:

$$\mathbf{x}_i = \mathbf{m} + \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_{d'} \mathbf{v}_{d'} + \dots + \alpha_d \mathbf{v}_d$$

Projection:

$$\mathbf{x}_i' = \mathbf{m} + \underbrace{\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_{d'} \mathbf{v}_{d'}}_{+ 0 \cdot \mathbf{v}_{d'+1} + \dots + 0 \cdot \mathbf{v}_d}$$

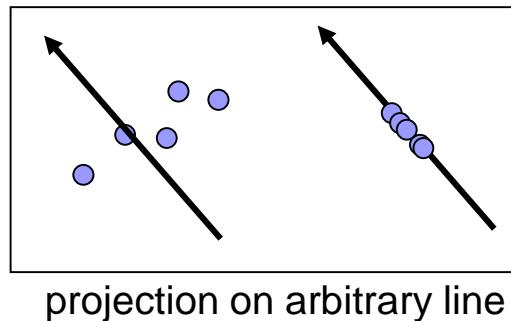
# Optimality of approximation

- The approximation is optimal in **least-squares sense**. It gives the minimal of:

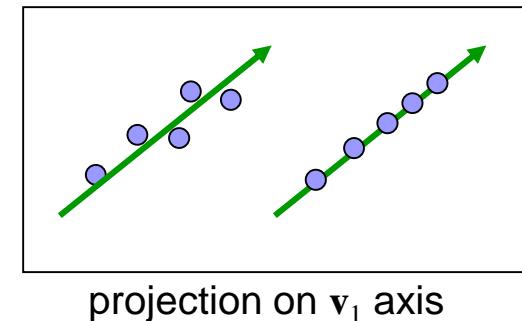
$$\sum_{k=1}^n \|\mathbf{x}_k - \mathbf{x}'_k\|^2$$

- The projected points have maximal variance.

Original set



projection on arbitrary line



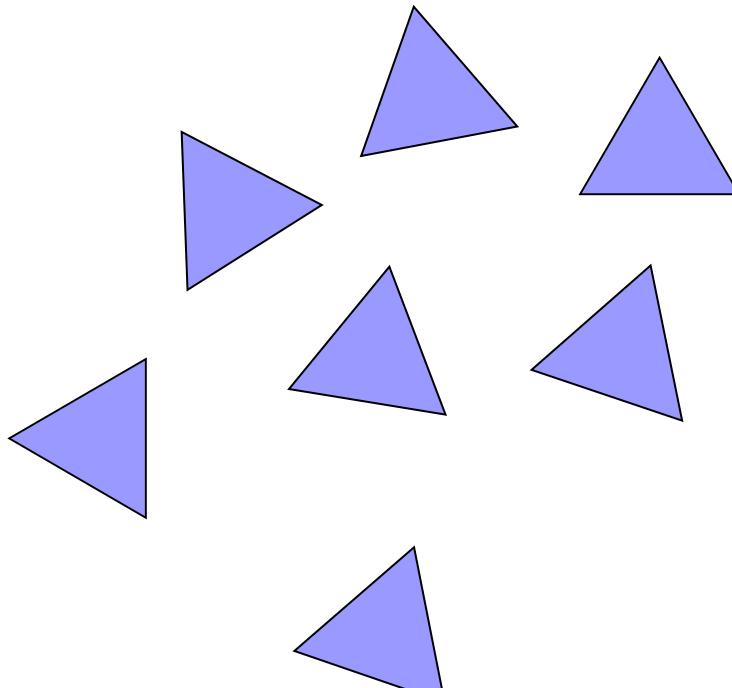
projection on v<sub>1</sub> axis

# PCA on Point Sets

$$S = \begin{pmatrix} y_1^1 & y_2^1 & \vdots & y_n^1 \\ y_1^2 & y_2^2 & & y_n^2 \\ \vdots & \vdots & & \vdots \\ y_1^d & y_2^d & \vdots & y_n^d \end{pmatrix} \begin{pmatrix} y_1^1 & y_1^2 & \vdots & y_1^d \\ y_2^1 & y_2^2 & \vdots & y_2^d \\ \vdots & \vdots & & \vdots \\ y_n^1 & y_n^2 & \vdots & y_n^d \end{pmatrix}$$

[demo](#)

# PCA on Geometric Primitives?



$$C_i = \int_{s_i} x x^T dx$$

Coordinate relative to center of mass