Towards Understanding More Complex Data

Graph Laplacian on Singular Manifolds

Yusu Wang

Joint work with M. Belkin, Q. Que, and X. Zhou

Computer Science and Engineering Dept.
The Ohio State University

Manifold Assumption

- Understanding data / spaces
  - Information estimation, data processing (clustering, semi-supervised learning, etc)

- Manifold assumption

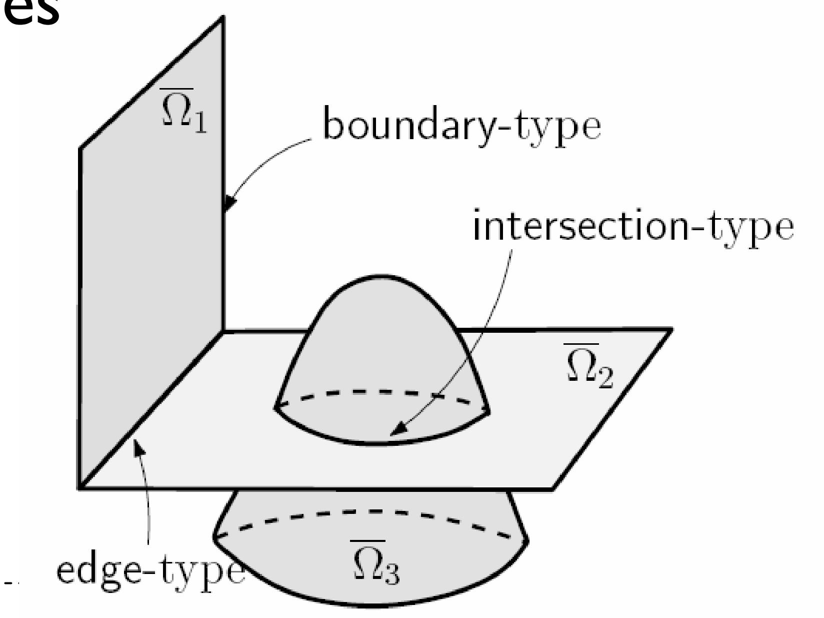
Manifold Assumption:

•Reasonable model for non-linear data

•Provides nice structures and properties that algorithms can leverage

Relax Manifold Assumption

- To model more complex data
  - Aim to relax manifold assumption
  - Still keep some nice properties that manifold assumption can offer
- Allow three types of "singularities"
  - Boundary-type
  - Intersection-type
  - Edge-type



Singularities

- Boundary
  - Configuration space may be constrained / limited

- Intersection-type
  - Two classes of data may contain similar instances

Goal:
Study how singularities may influence learning / information retrieval algorithms and how they can be learned from data.

This Talk

- Aim to study singular manifolds through the lens of Gaussian-weighted graph Laplacian
  - Laplacian-based methods is a widely used class of techniques used for recovering geometric properties of data.
  - There is a fairly good theoretical understanding of properties of Laplacian when data is sampled from a smooth manifold.
- Goal of this talk:
  - Behavior of graph Laplacian for data sampled from a singular manifold
    - when singularities are present

Some Related Work

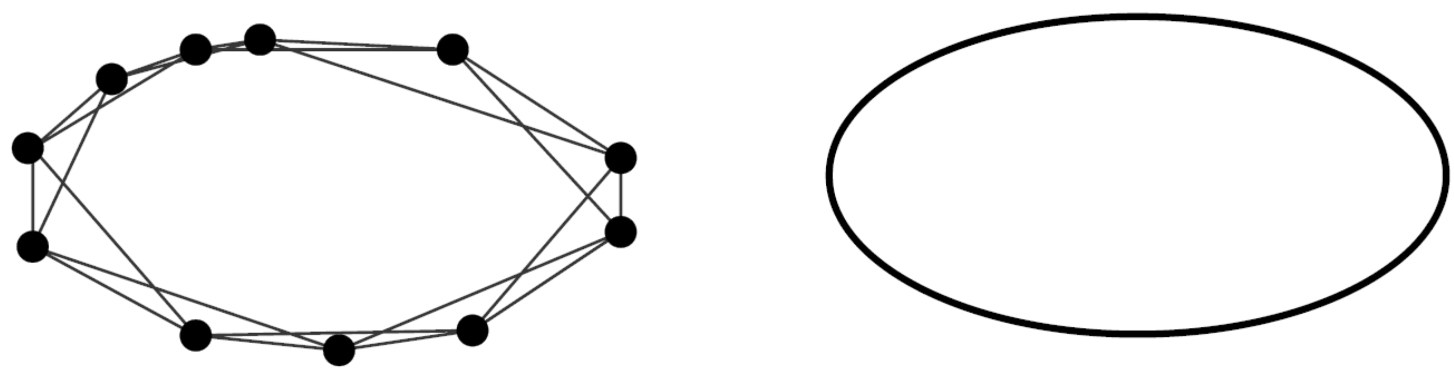
- Sampling theory for compact domains
  - [Chazal, Cohen-Steiner, Lieutier 09], [Chazal,Oudot 08], [Cheng,Dey,Ramos 07], …
- Learning collection of linear sub-spaces
  - [Vidal, Ma, Sastry 05], [Chen, Lerman 09], …
- Learning stratified spaces
  - [Bendich, Cohen-Steiner, Edelsbrunner, Harer, Morozov 07], [Bendich, Wang, Mukherjee 12], …

This Talk

- Introduction
- Graph / Functional Laplacian
- Behavior of Functional Laplacian on / around singularities
  - Boundary-type
  - Edge-type
  - Intersection-type
- Discussion and implications

Extract Manifold

- Data from a hidden smooth manifold
- Construct a graph that describes the manifold
  - Properties of graph reflect those of manifold



What Property?

- Gaussian weighted graph Laplacian
- *n* data points: $P = \{p_1, p_2, \ldots, p_n\}$
- L_P^t: *n x n* matrix

$$\mathsf{L}_P^t[i][j] = \begin{cases} -\frac{1}{n} \cdot \frac{1}{(4\pi t)^{k/2} t} e^{-\frac{\|p_i - p_j\|^2}{4t}}, & \text{if } i \neq j \\ \frac{1}{n} \cdot \frac{1}{(4\pi t)^{k/2} t} \sum_{l \neq i, l \in [1,n]} e^{-\frac{\|p_i - p_l\|^2}{4t}}, & \text{if } i == j \end{cases}$$

- When it performs on a function (*n*-vector) *f*:

Graph Laplacian is a light-weight structure (depending only proximity graph), suitable for high dimensional data analysis.

Laplace-Beltrami Operator

Nice properties of manifold Laplacian

- Reflect manifold geometry
- Eigenfunctions form basis for functions on manifold
- Relation to heat operator
- …

Applications

- Clustering, semi-supervised learning
- Data denoising
- Data representation
- Graphics, mesh smoothing, optimization

Functional Laplacian

- Laplace-Beltrami operator Δ_M is useful
- Gaussian weighted graph Laplacian L_P^t approximates Δ_M
  - For points *P* uniformly randomly sampled from manifold *M*.
  - L_P^t pointwise-converges to Δ_M *[BN05]*
  - Spectral convergence *[BN08]*
- Connection made through functional Laplacian

$$L_t f(x) = \frac{1}{t(4\pi t)^{d/2}} \int_M e^{-\frac{\|y-x\|^2}{4t}} (f(y) - f(x)) d\nu(y)$$

  - can be considered the limit of L_P^t as the size of *P* goes to ∞.

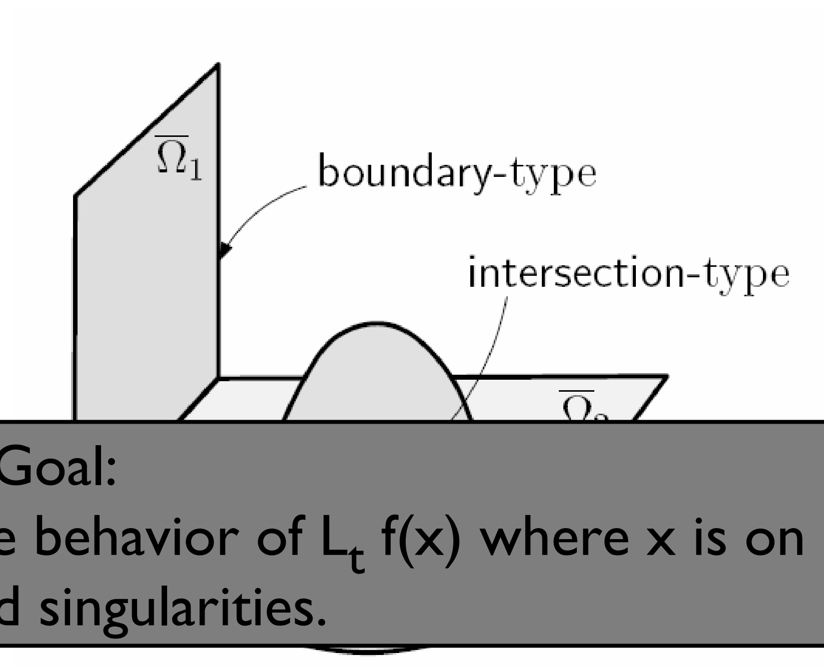
$$\boxed{\mathsf{L}_P^t f(p_i) = \frac{1}{n} \cdot \frac{1}{(4\pi t)^{k/2} t} \sum_{j=1}^{n} e^{-\frac{\|p_i - p_j\|^2}{4t}} (f(p_i) - f(p_j))}$$

This Talk

- Introduction
- Graph / Functional Laplacian
- Behavior of Functional Laplacian on / around singularities
  - Boundary-type
  - Edge-type
  - Intersection-type
- Discussion and implications

Singular Manifold

- A singular manifold Ω is a collection of smooth manifolds with boundaries $\Omega_1, \Omega_2, \ldots, \Omega_m$
- A point x is
  - Boundary type
    - If $x \in \partial\Omega_i$
  - Intersection type
    - If $x \in \Omega^o{}_i \cap \Omega^o{}_j$
  - Edge type



Goal:
Given a function f, analyze the behavior of $L_t f(x)$ where x is on or around singularities.

Laplacian at a Regular Point

- At a regular point (an interior point of a manifold) *x*

$$\boxed{L_t f(x) = \frac{1}{t(4\pi t)^{d/2}} \int_M e^{-\frac{\|y-x\|^2}{4t}} (f(y) - f(x)) d\nu(y)}$$

- By Taylor expansion:
  - $f(y) \approx f(x) + (y-x)^T \nabla f(x) + (y-x)^T H (y-x)$

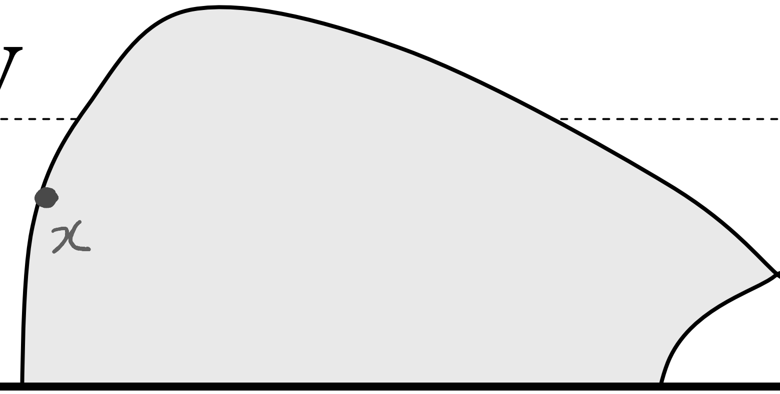
$$\begin{aligned} L_t f(x) &\approx \frac{1}{t} \int_M K_t(x,y)((y-x)^T \nabla f(x) + (y-x)^T H(y-x)) dy \\ &= \frac{1}{t} \int_M K_t(x,y)(y-x)^T H(y-x) dy \end{aligned}$$

= C tr(H)

$$\Rightarrow \quad \boxed{L_t f(x) = C \cdot \Delta f(x) + o(1)}$$

Laplacian at Boundary

- At a boundary point x



$$
\begin{aligned}
L_t f(x) &= \frac{1}{t}\int_M K_t(x,y)(f(y)-f(x))dy \\
&\approx \frac{1}{t}\int_M K_t(x,y)((y-x)^T\nabla f(x) + (y-x)^T H(y-x))dy \\
&= \boxed{\frac{1}{t}\int_M K_t(x,y)(y-x)^T\nabla f(x)dy} \\
&\quad + \frac{1}{t}\int_M K_t(x,y)(y-x)^T H(y-x))dy
\end{aligned}
$$

O(1)

Intuitive Illustration

- First term:

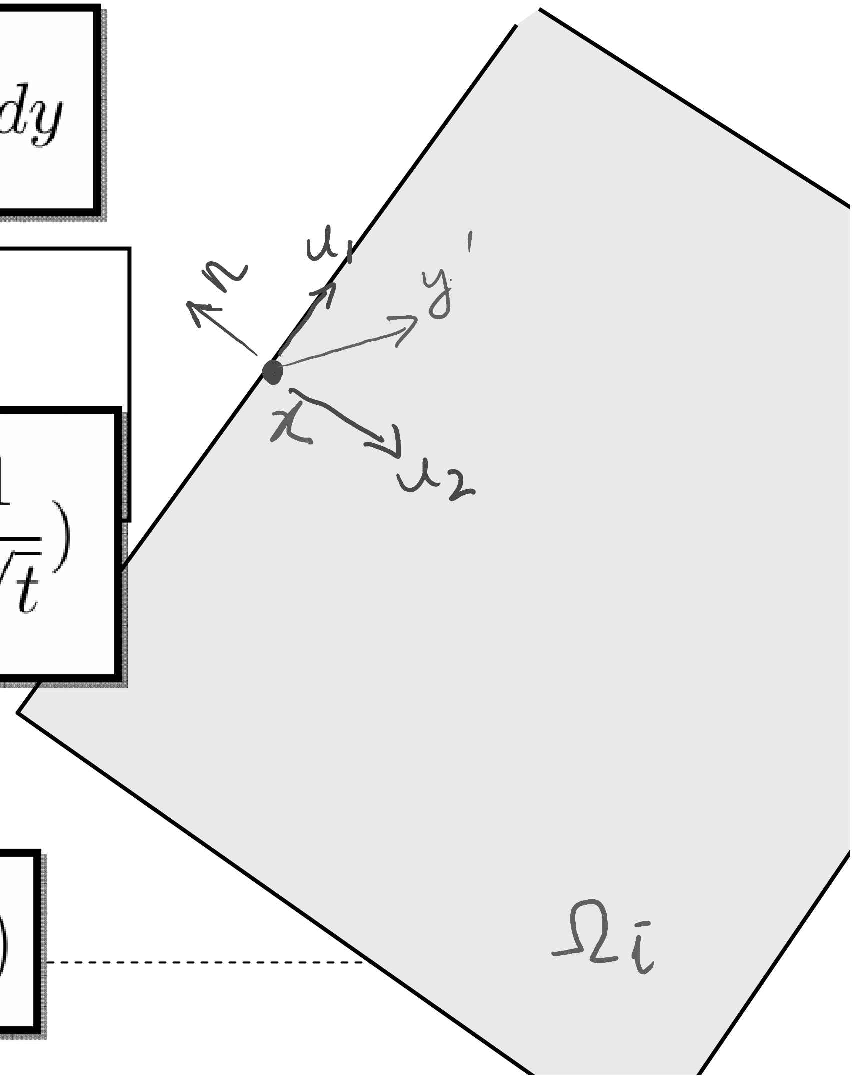
$$\frac{1}{t}\int_M K_t(x,y)(y-x)^T\nabla f(x)dy$$

For a boundary point x

$$L_t f(x) = -\frac{1}{\sqrt{t}}C_1\partial_{\mathbf{n}}f(x) + o(\frac{1}{\sqrt{t}})$$

For a regular point x

$$L_t f(x) = C\cdot\Delta f(x) + o(1)$$



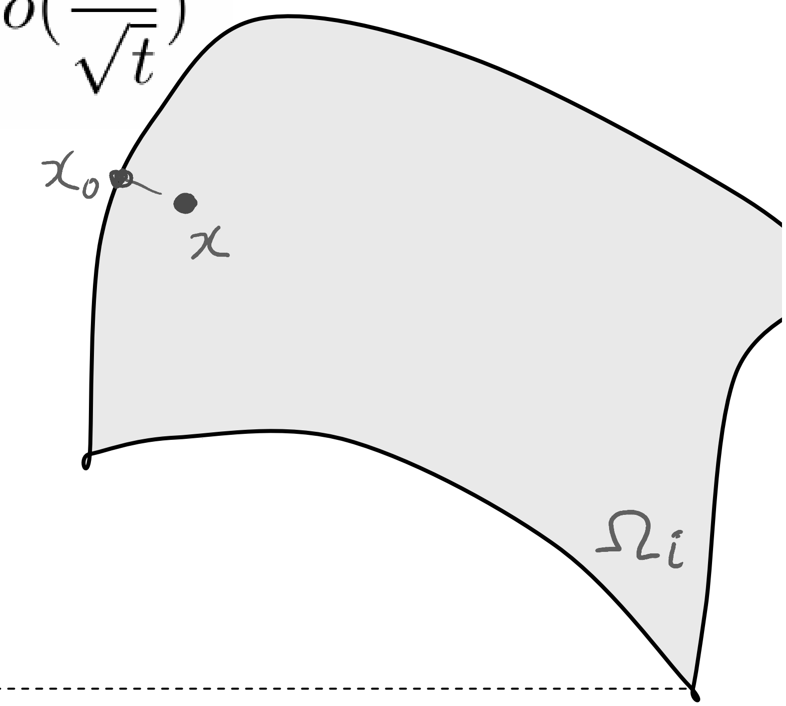
Laplacian Around Boundary

- For a point *x* near boundary
- Let x_0 be nearest neighbor of *x* along the boundary
  - Set $\| x - x_0 \| = r\sqrt{t}$

$$L_t f(x) = -\frac{1}{\sqrt{t}} e^{-r^2} C_1 \partial_{\mathbf{n}} f(x_0) + o(\frac{1}{\sqrt{t}})$$

As x moves away from the boundary, the boundary effect decreases rapidly.

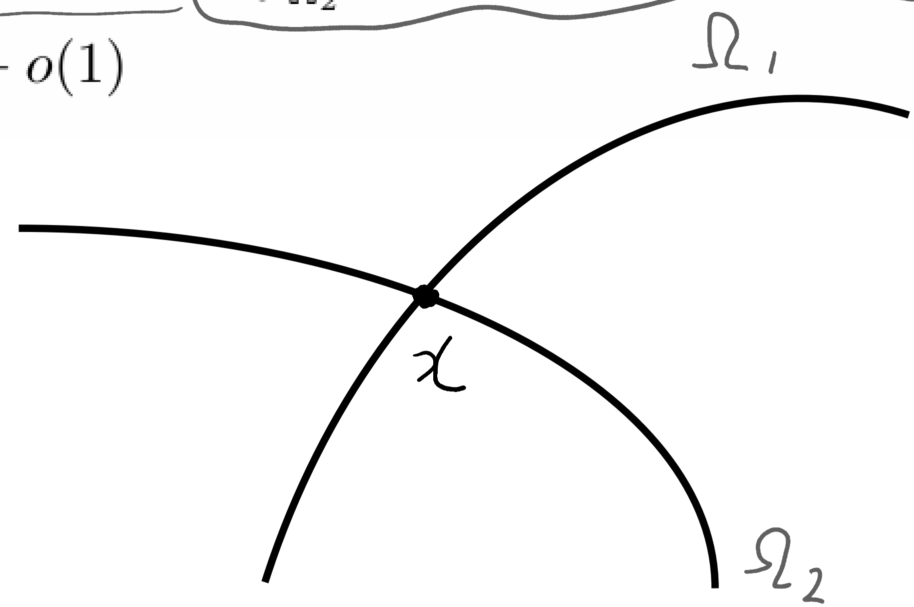
Points roughly within $\sqrt{t}$ distance to boundary have boundary effect.



Laplacian at Intersection Singularity

- For a point *x* on intersection singularity

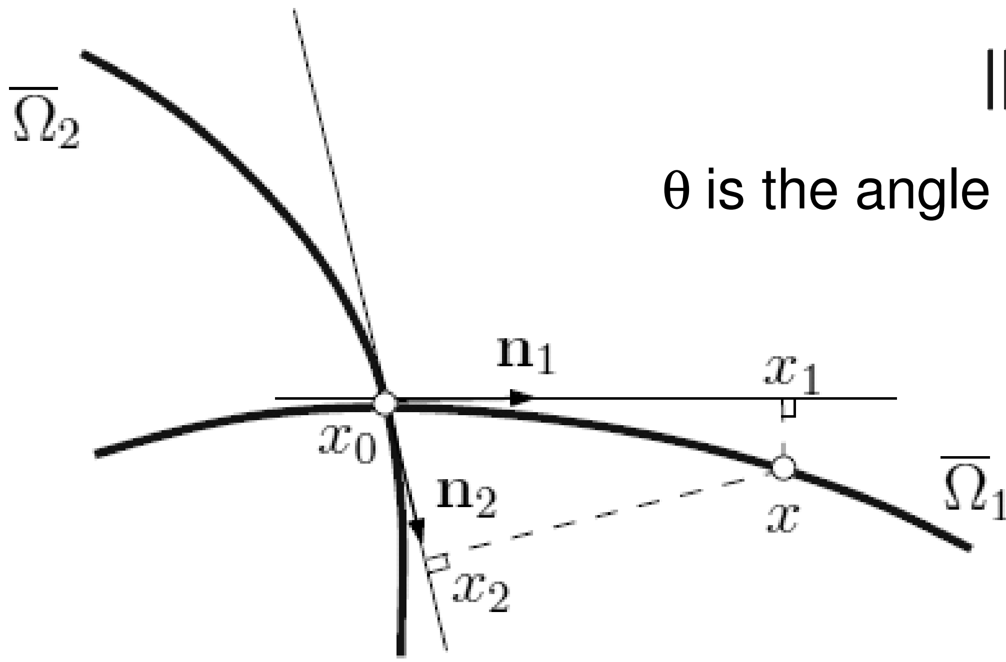
$$
\begin{aligned}
L_t f(x) &= \frac{1}{t}\int_{\Omega_1\cup\Omega_2} K_t(x,y)((f(y)-f(x))dy \\
&= \frac{1}{t}\int_{\Omega_1} K_t(x,y)((f(y)-f(x))dy + \frac{1}{t}\int_{\Omega_2} K_t(x,y)((f(y)-f(x))dy \\
&= C\Delta_{\Omega_1} f(x) + C\Delta_{\Omega_2} f(x) + o(1)
\end{aligned}
$$



Laplacian Around Intersection

- For a point x around intersection singularities

$$L_t f(x) = \boxed{\frac{1}{\sqrt{t}}}\boxed{re^{-r^2 \sin^2 \theta}} C_2(\boxed{\partial_{\mathbf{n}_1} f_1}(x_0) + \cos\theta \cdot \boxed{\partial_{\mathbf{n}_2} f_2}(x_0)) + o(\frac{1}{\sqrt{t}})$$



$\| x - x_0 \| = r\sqrt{t}$

θ is the angle between n_1 and n_2

Laplacian On / Near Edge-Singularities

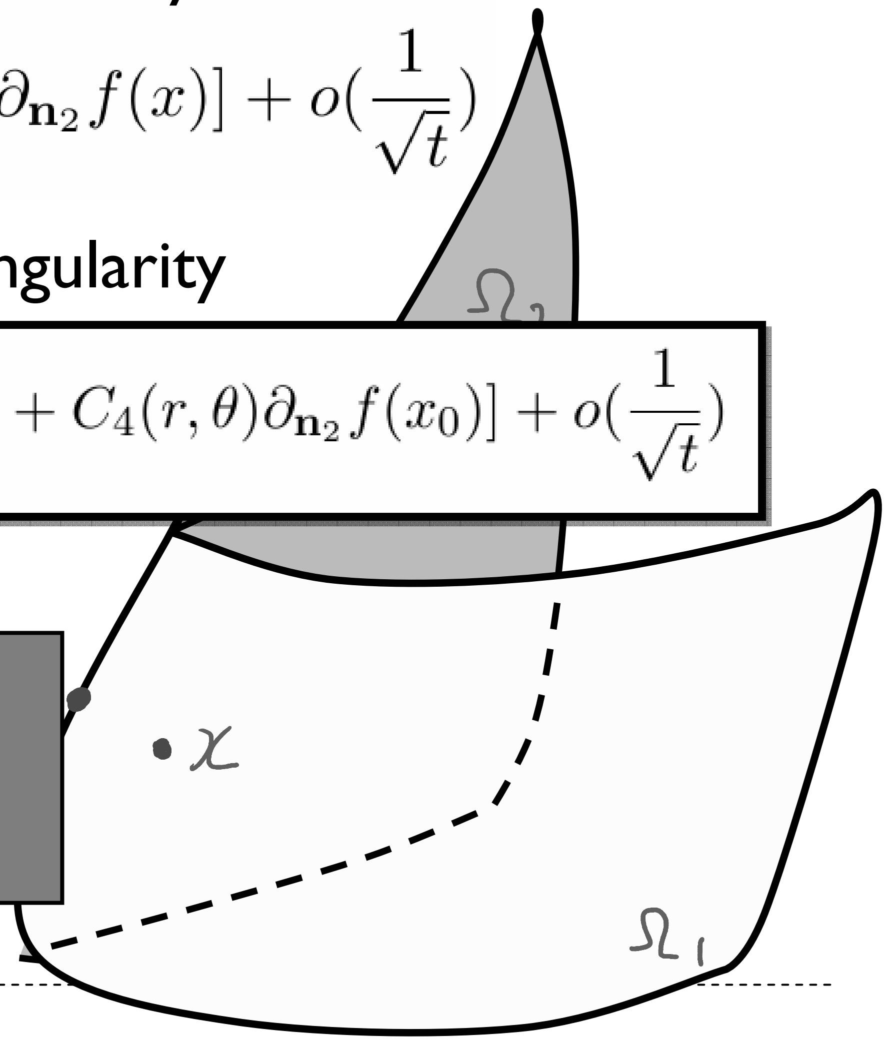
- For a point x on a "glued" boundary of two manifolds

$$L_t f(x) = -\frac{1}{\sqrt{t}} C_1[\partial_{\mathbf{n}_1} f(x) + \partial_{\mathbf{n}_2} f(x)] + o(\frac{1}{\sqrt{t}})$$

- For a point x near an edge singularity

$$\boxed{L_t f(x) = -\frac{1}{\sqrt{t}}[C_3(r,\theta)\partial_{\mathbf{n}_1} f(x_0) + C_4(r,\theta)\partial_{\mathbf{n}_2} f(x_0)] + o(\frac{1}{\sqrt{t}})}$$

Have both boundary effect and effects from points of the other manifold.



This Talk

- Introduction
- Graph / Functional Laplacian
- Behavior of Functional Laplacian on / around singularity
  - Boundary-type
  - Edge-type
  - Intersection-type
- Discussion and implications

Key Point I

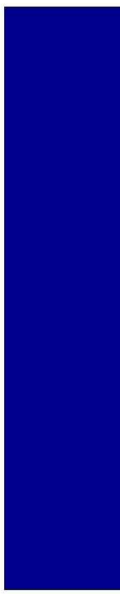
- Graph Laplacian around all three types of singularities
  - Bnd: $L_t f(x) = -\frac{1}{\sqrt{t}} e^{-r^2} C_1 \partial_{\mathbf{n}} f(x_0) + o(\frac{1}{\sqrt{t}})$
  - Intersec: $L_t f(x) = \frac{1}{\sqrt{t}} r e^{-r^2 \sin^2\theta} C_2(\partial_{\mathbf{n}_1} f_1(x_0) + \cos\theta \cdot \partial_{\mathbf{n}_2} f_2(x_0)) + o(\frac{1}{\sqrt{t}})$
  - Edge: $L_t f(x) = -\frac{1}{\sqrt{t}} [C_3(r,\theta)\partial_{\mathbf{n}_1} f(x_0) + C_4(r,\theta)\partial_{\mathbf{n}_2} f(x_0)] + o(\frac{1}{\sqrt{t}})$

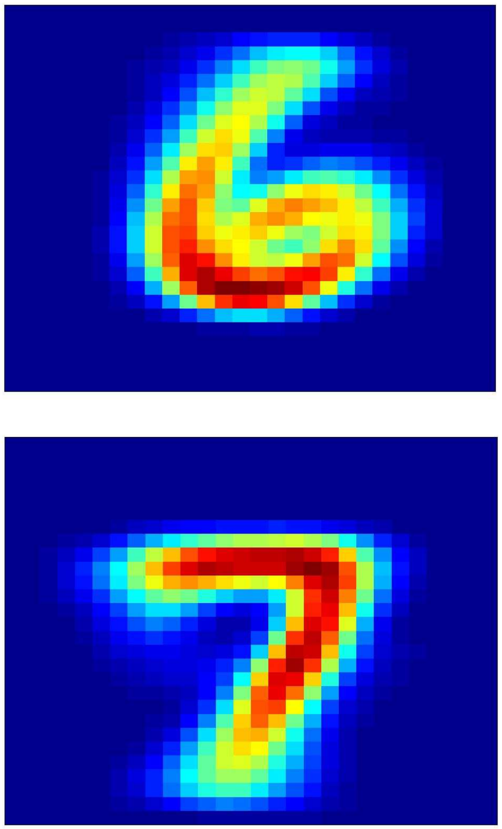
- L_t *f(x)* at points around singularities have significantly different behavior

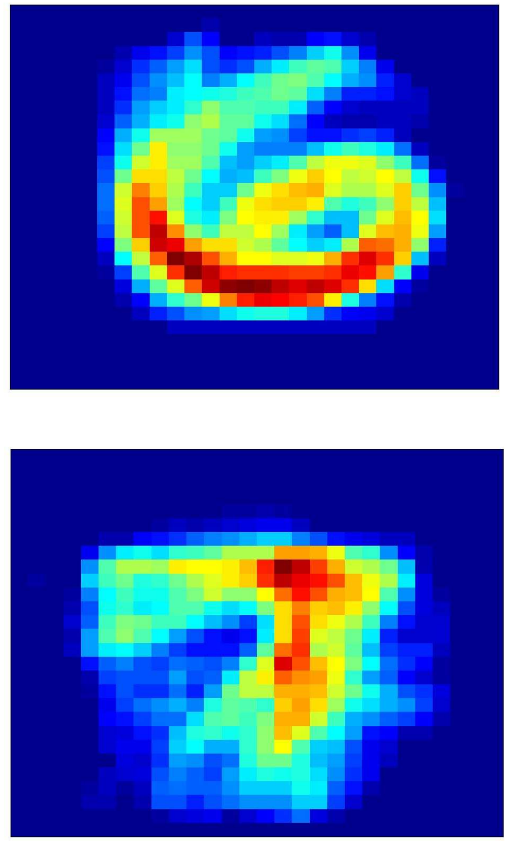
This suggests potentially identifying singularities by finding points with high $L_t f(x)$ values.

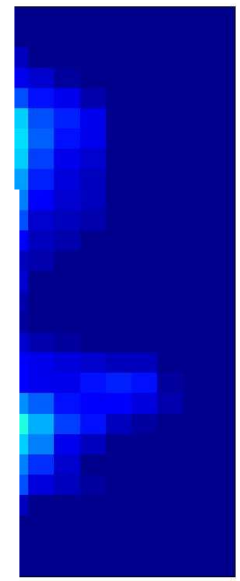
Examples

- MNIST digit images
  - Ea ... ure space
  - Fu







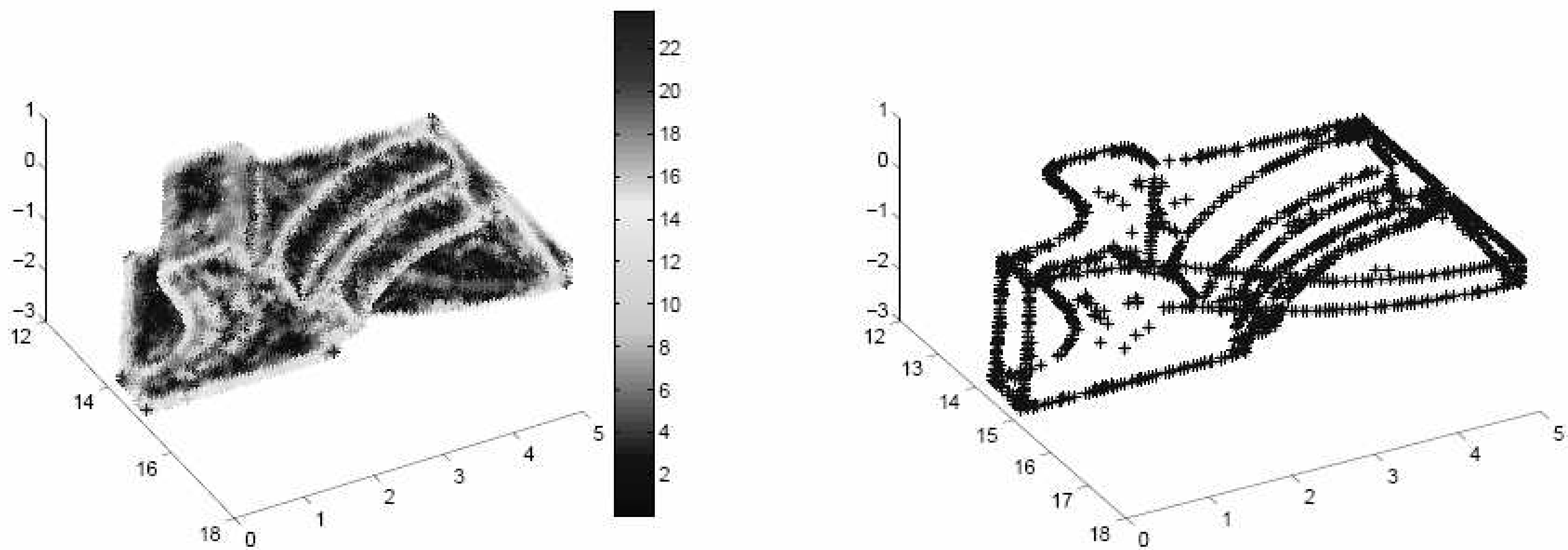


Digit ... average image

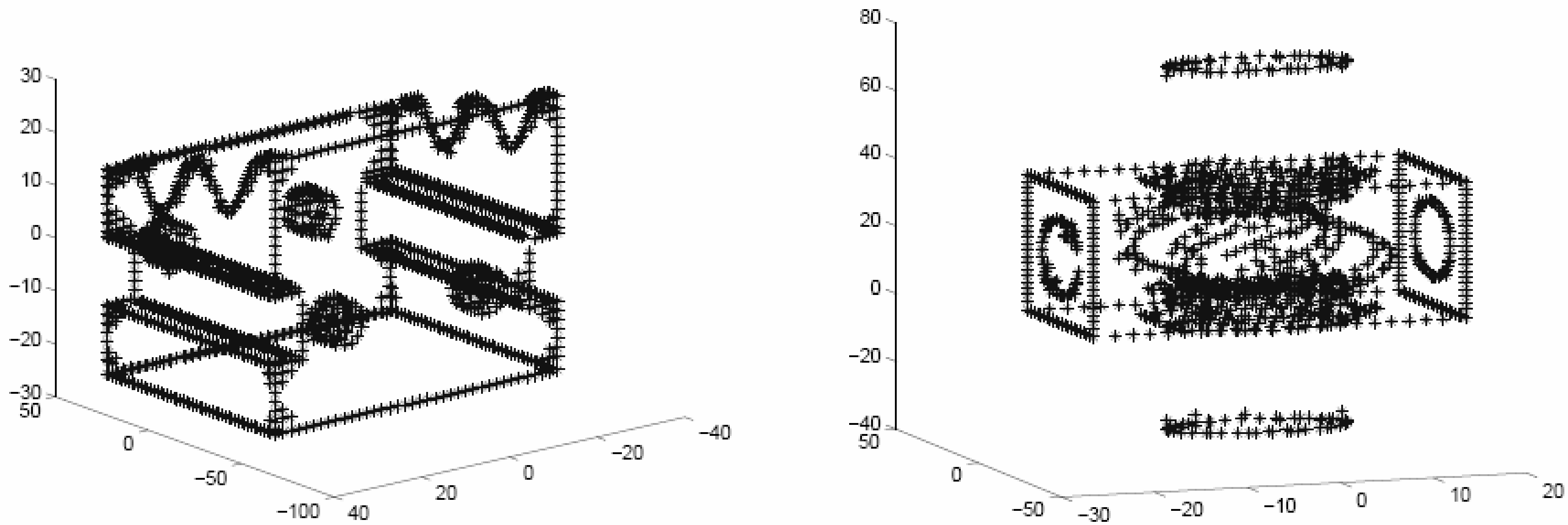
singularity image ... larity image

Examples

- Sharp feature curves identification from point clouds
- Apply L_t to coordinate functions

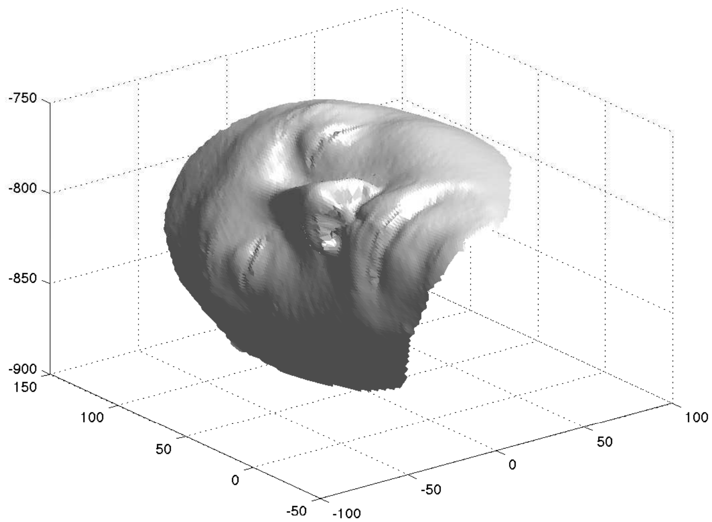


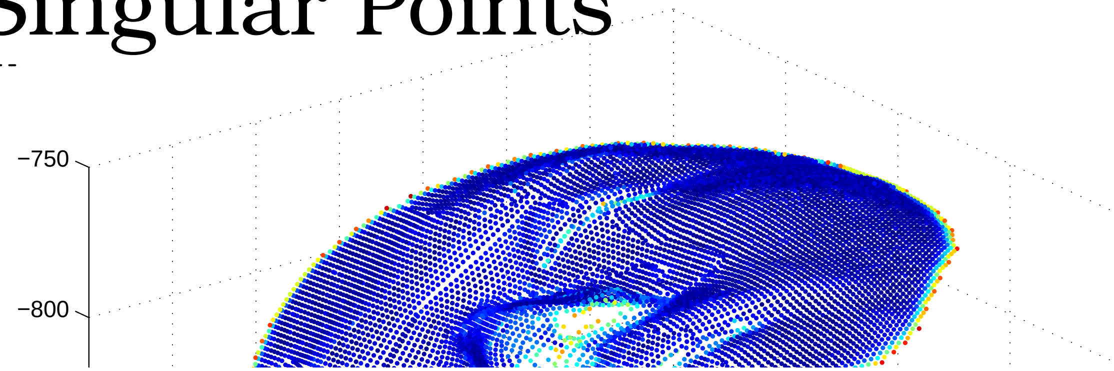
(a) Norm of V_t on model of fandisk. (b) Points with large norm.

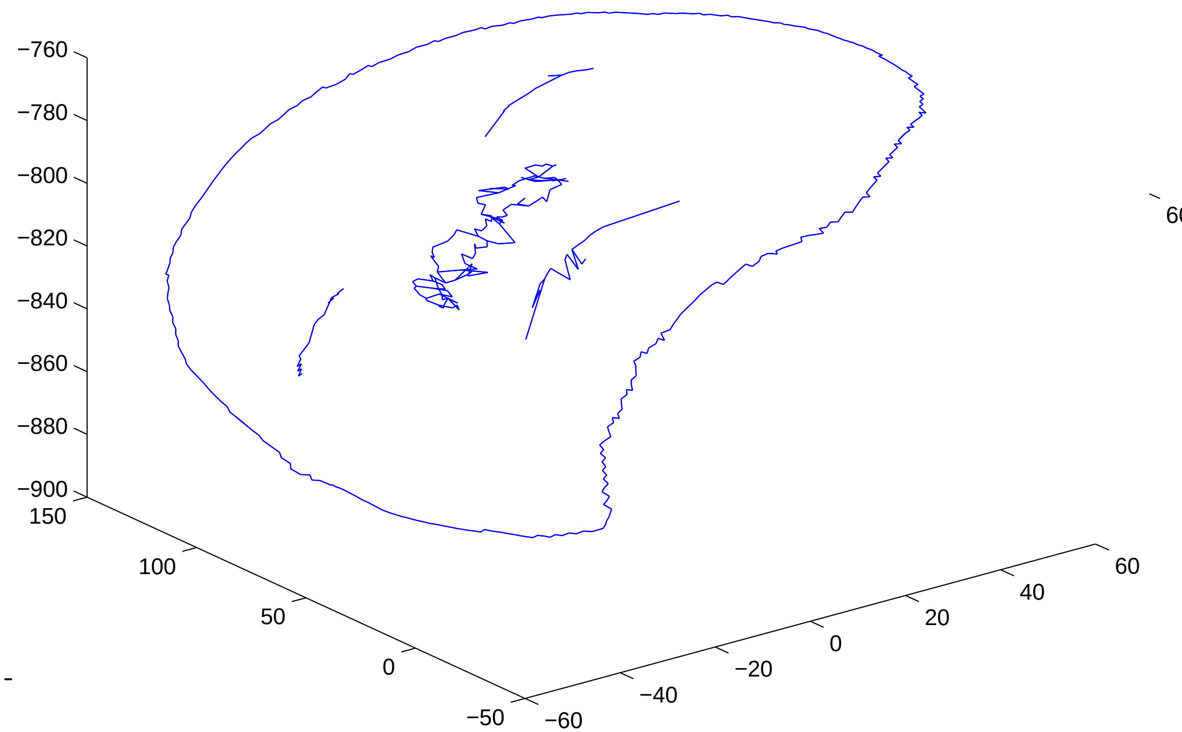

30
20
10
0
−10
−20
−30
50
0
−50
−100
40
20
0
−20
−40
80
60
40
20
0
−20
−40
50
0
−50
−30
−20
−10
0
10
20

(a) (b)

Reeb Graph + Singular Points

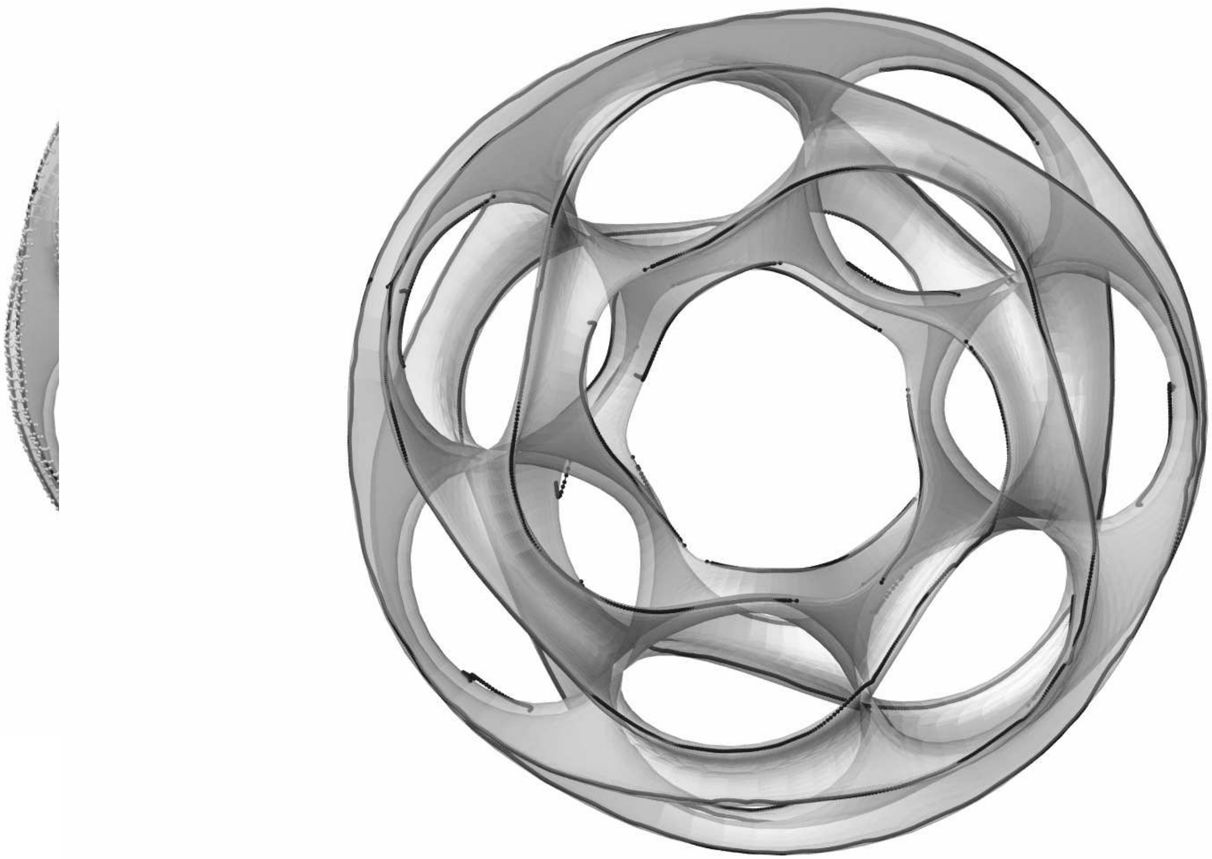






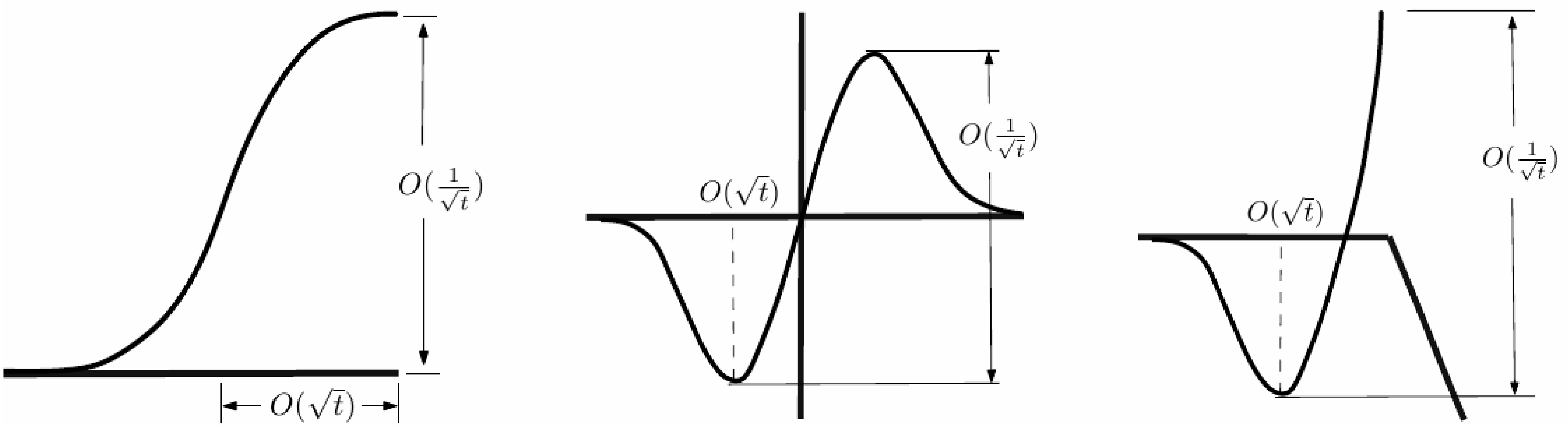
Obtain feature curve

Reeb Graph + Singular Points



Key Point II

- Around different types of singularities, scaling behaviors are different.

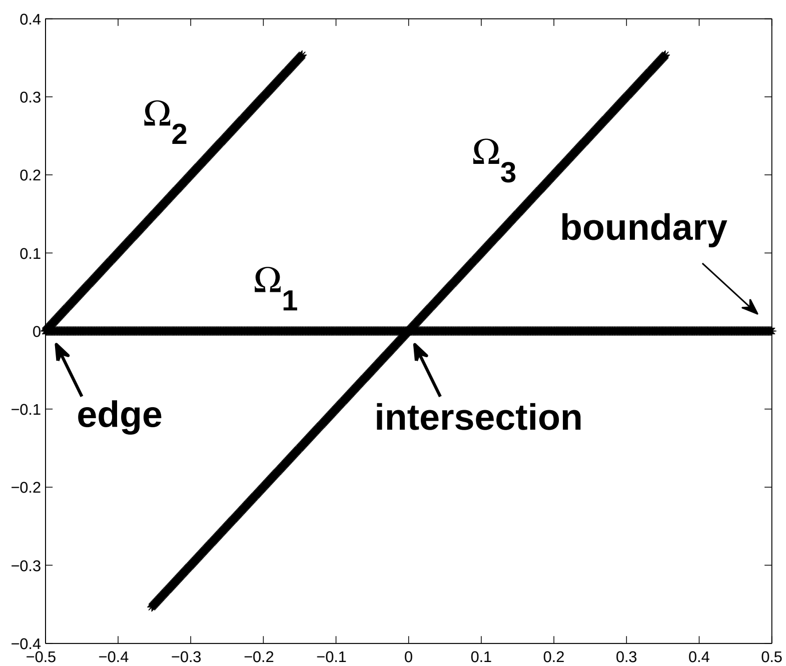


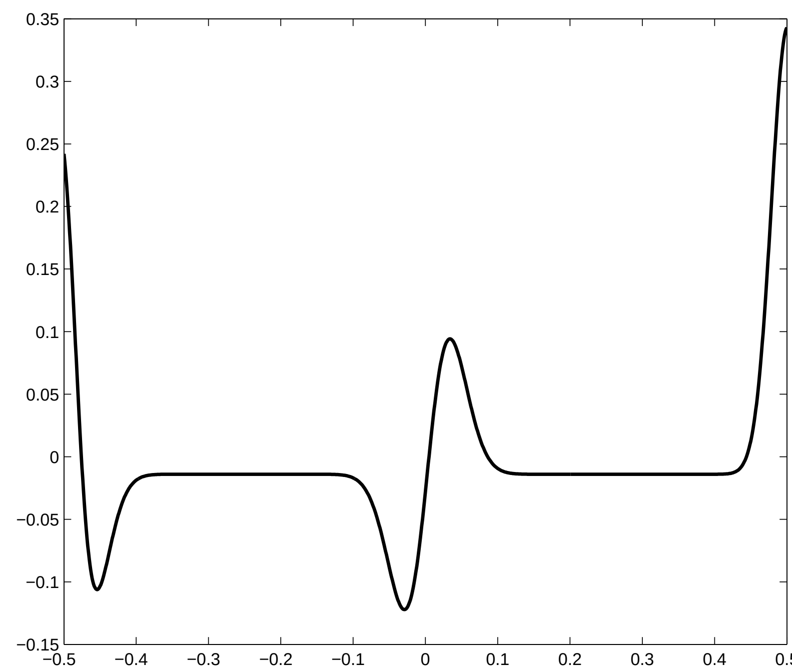
Bour ... lge

These local behaviors can potentially be used to distinguish different type of singularities

$\phi(r) = \ldots \sqrt{t} \ldots \sqrt{t} \ldots \sqrt{t} \ldots {}^{-r^2} + re^{-r^2})$

Examples



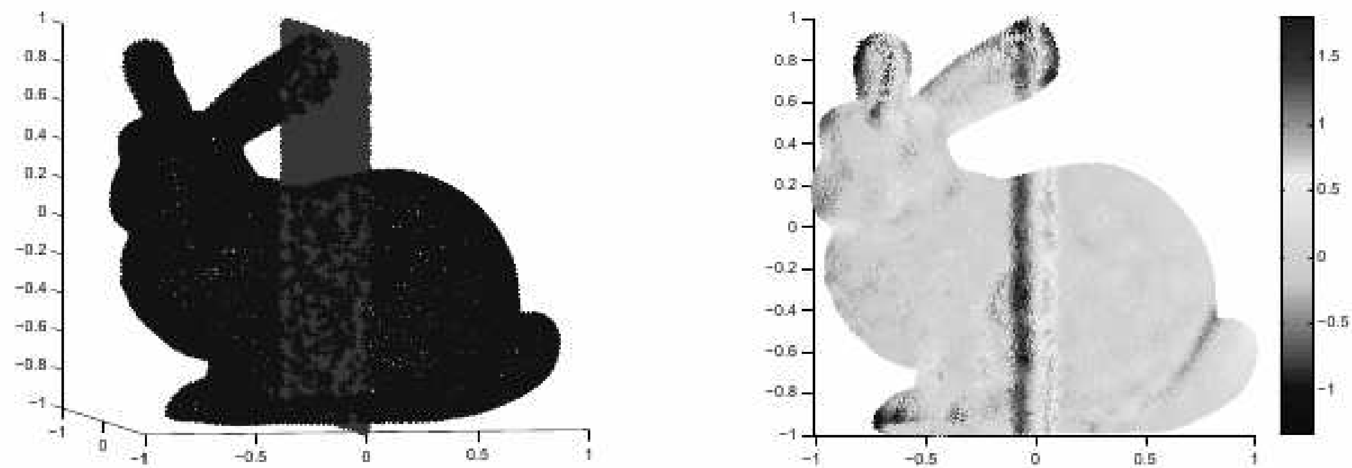


Data Domain

$L_t f$ on Ω_1, where

$$f(x_1, x_2) = (x_1+0.2)^2 + x_2^2$$

Examples



(a) Data set

(b) Values of $L_{n,t}f$ on Bunny when $f = y$.

Disclaimer: No animals were harmed during the making of this slide.

Key Point III

- Can singularity points be simply ignored?
  - After all, these points are of measure zero

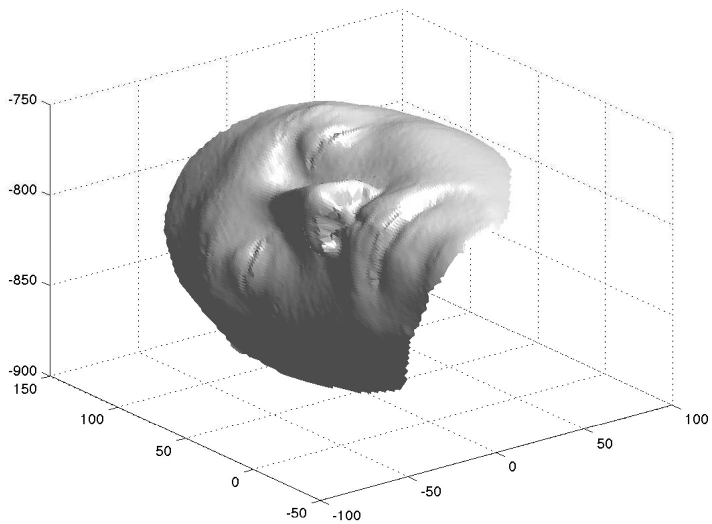
- No. At least not for singularity of codimension-one
  - Roughly, the total effect is $O(\frac{1}{\sqrt{t}}) \cdot O(\sqrt{t}) = O(1)$
  - As t tends to 0, this effect does not vanish

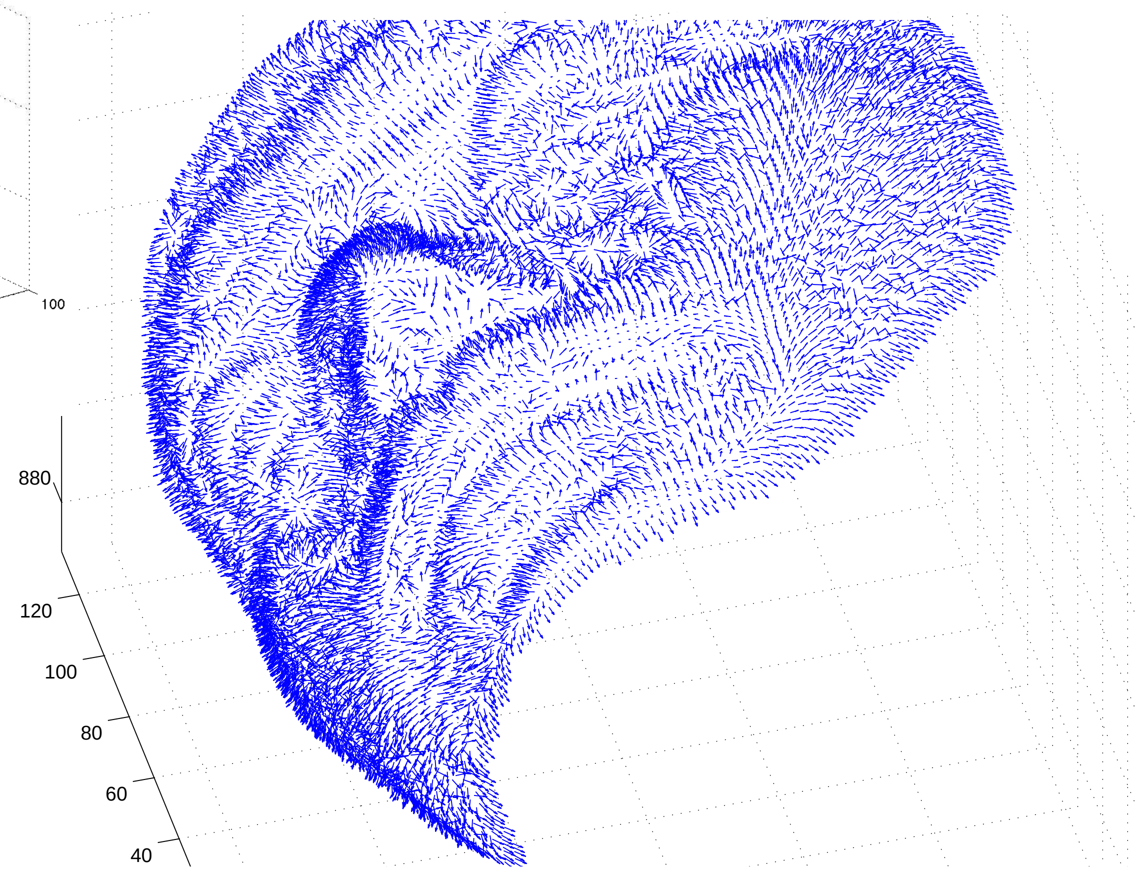
Key Point IV

- Recall for a boundary point x

$$L_t f(x) = -\frac{1}{\sqrt{t}} C_1 \partial_{\mathbf{n}} f(x) + o(\frac{1}{\sqrt{t}})$$

  - It can be used to compute outward-normal at boundary
  - It can also be used to compute the partial derivative of a function along outward-normal at boundary.


-750
-800
-850
-900
150
100
50
0
-50
-100
-50
0
50
100


880
120
100
80
60
40

Eigen-functions

- Consider any point x at boundary singularity
- Let ϕ_t be an eigenfunction for L_t .

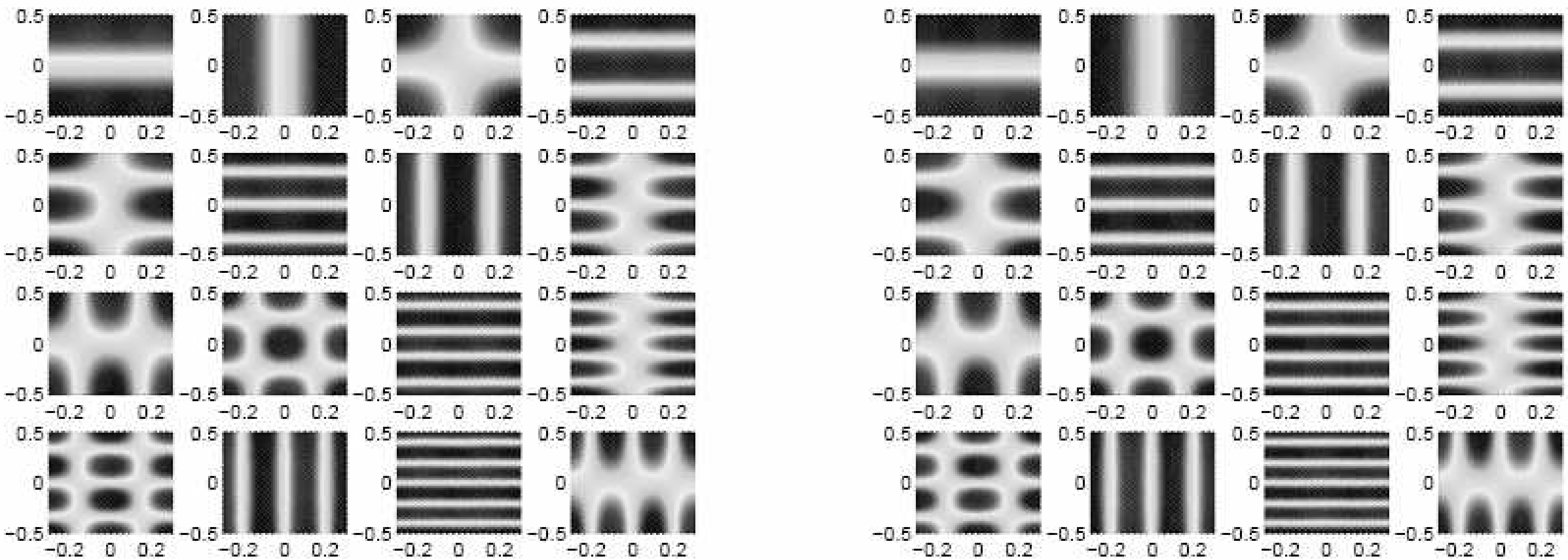
$$-\frac{1}{\sqrt{t}} \cdot C \cdot \partial_{\mathbf{n}} \phi_t(x) \approx L_t \phi_t(x) = \lambda_t \phi_t(x) < \infty$$

- As t tends to 0, if the limit of ϕ_t exists,
  - it takes on Neumann boundary condition
- Initial numerical results seem to confirm this.

Eigen-functions

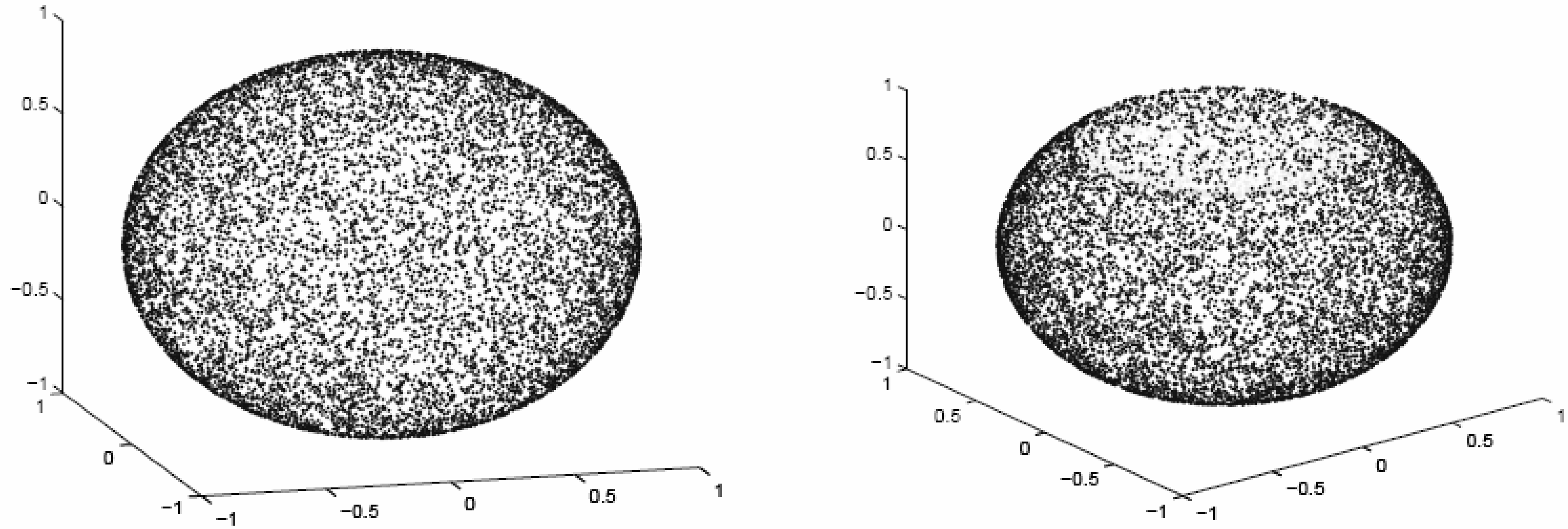
- Edge-type singularity
  - $\partial_{\mathbf{n}_1}\phi_t|_{\Omega_1}(x_0) = \partial_{\mathbf{n}_2}\phi_t|_{\Omega_2}(x_0)$
- Conjecture:
  - Eigenvalues and eigenfunctions of two isometric singular manifolds are the same.

Examples



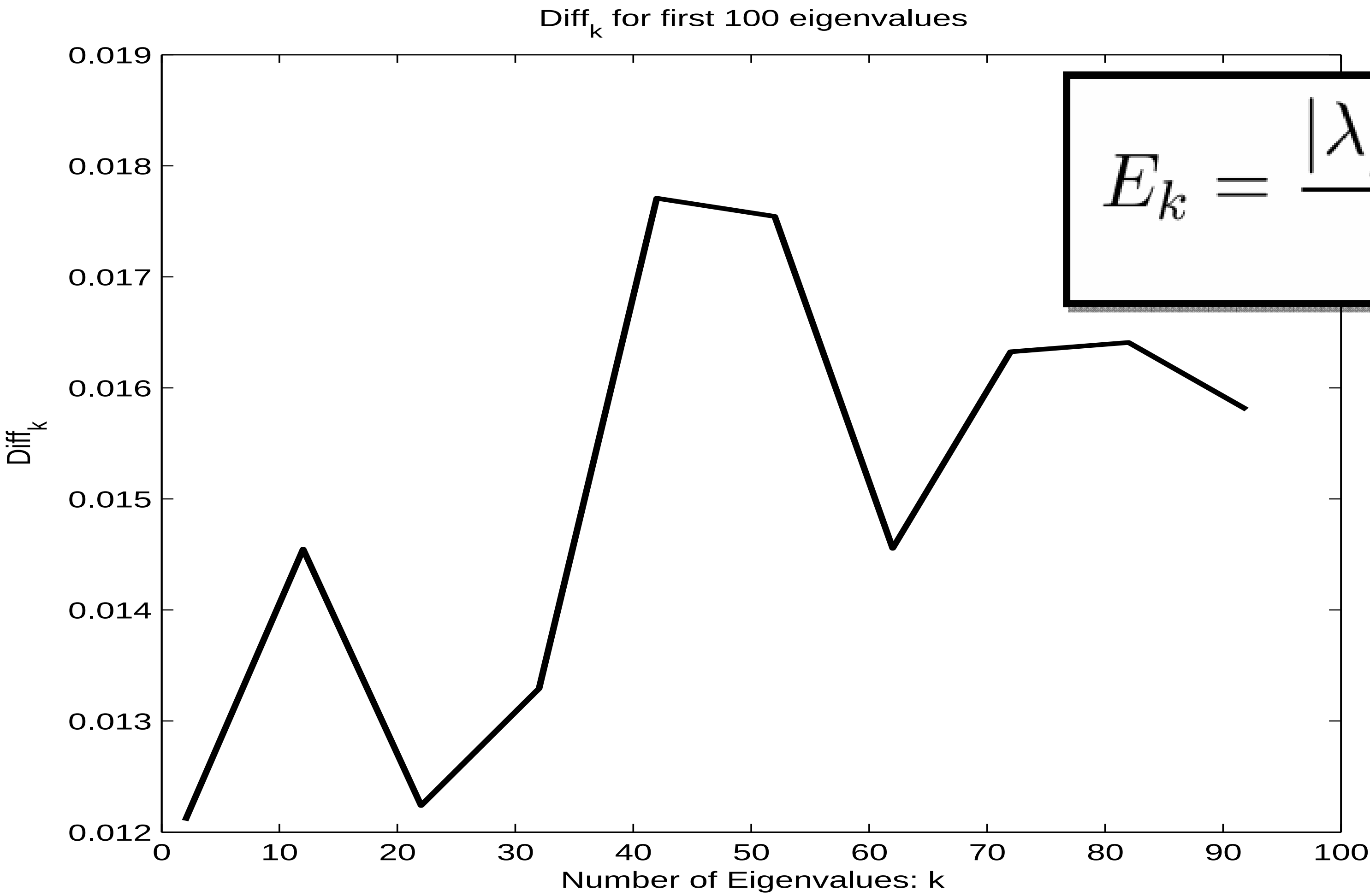
(a) Ω_1 (b) Ω_2

Figure 5: 2nd-17th eigenvectors of the graph Laplace matrices of the two manifolds Ω_1 and Ω_2



(a) S_1 (b) S_2

Figure 6: Unit sphere S_1 (left), and unit sphere with the top "sliced and flipped" S_2 (right).



$$E_k = \frac{|\lambda_k^1 - \lambda_k^2|}{|\lambda_k^1|}$$

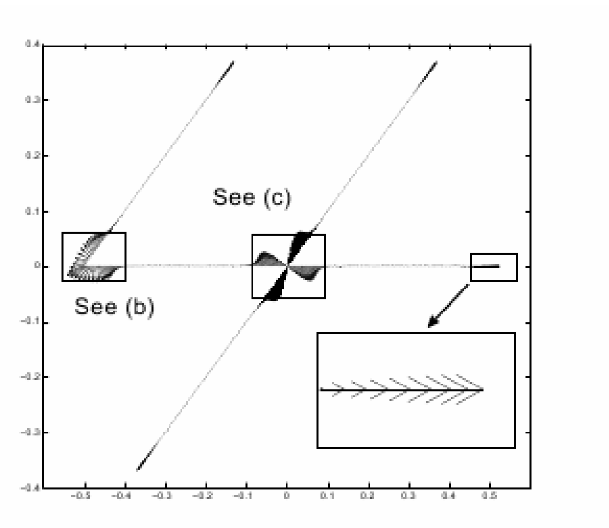
Summary and Other Discussions

- Initial study of the behavior of weighted Graph Laplacian on singular manifolds
  - Lightweight structure, geometry information it captures

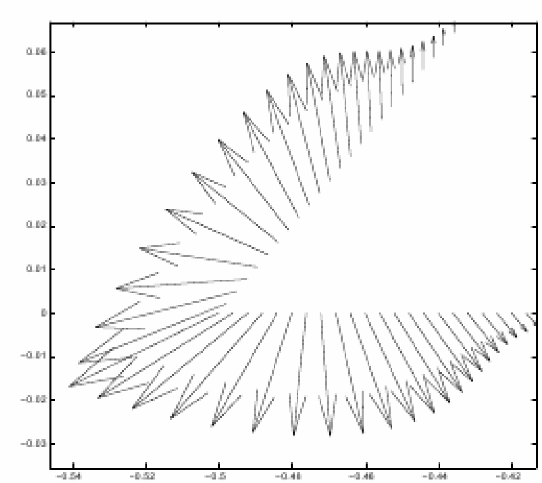
- Behavior of its eigenfunctions
- Potential applications:
  - Feature curve reconstruction for surface models from point samples
  - Better de-noising or classification algorithms ?
- Learning collection of linear subspaces ?
- Combining with stratification learning ?

Examples

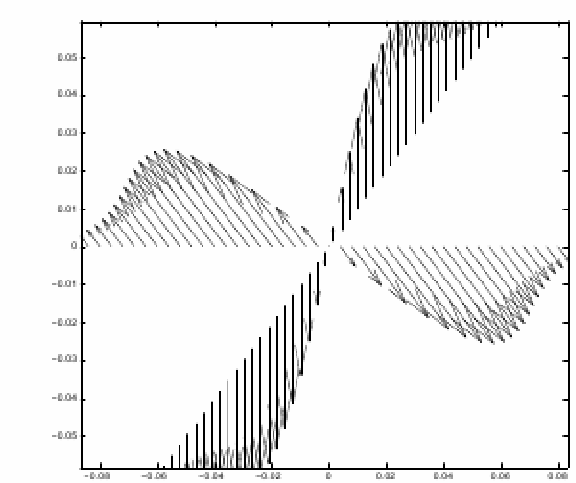
- Apply L_t to ambient space coordinate functions
  - induces a vector field



Input Data



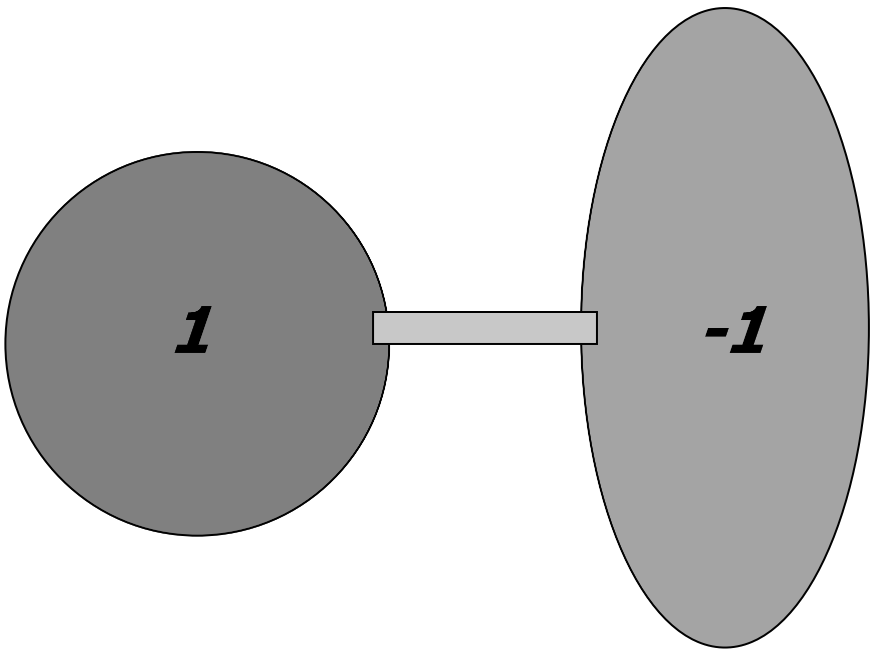
Edge-singularity



Intersectoin-singularity

Example: I

- Clustering:



- $\Delta f = 0$ if f is constant on each component
- Take Eigenfunction corresponding to 0 Eigenvalue
- Segmentation etc

Example II:

- Smoothing

- Eigenfunctions form a basis for functions on manifold
- Relation to Heat

- Levy: Laplace-beltrami eigenfunctions: Towards an algorithm that understands geometry. In *IEEE SMI, invited talk (2006)*
- Sorkine. Differential representations for mesh processing. *Computer Graphics Forum*, (2006).
- Zhang, Van Kaick, Dyer: Spectral mesh processing. *Computer Graphics Forum,* (2009).