Topological Persistence

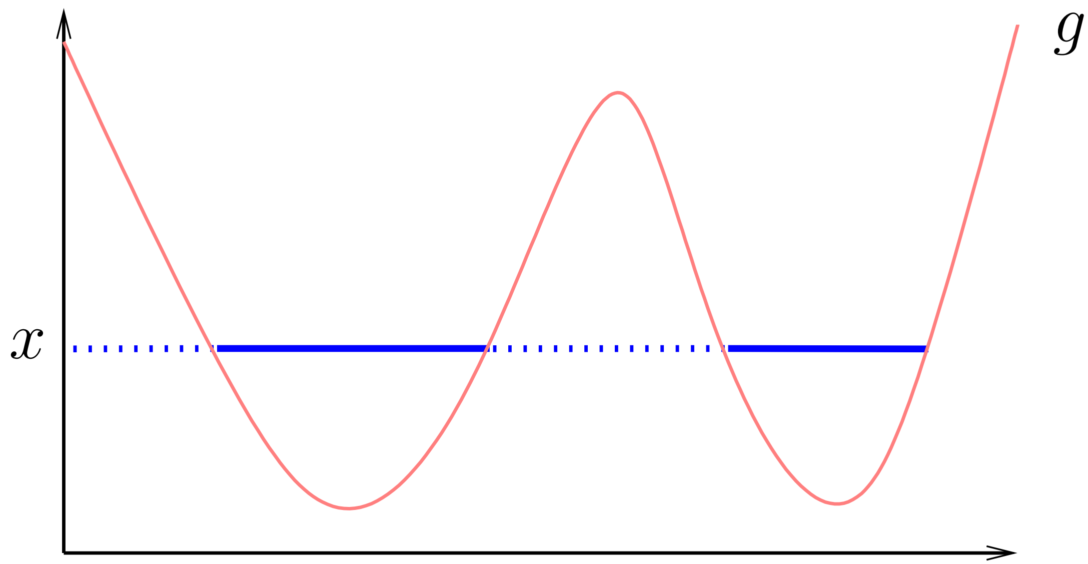
David Cohen-Steiner

INRIA
SOPHIA ANTIPOLIS

Introduction

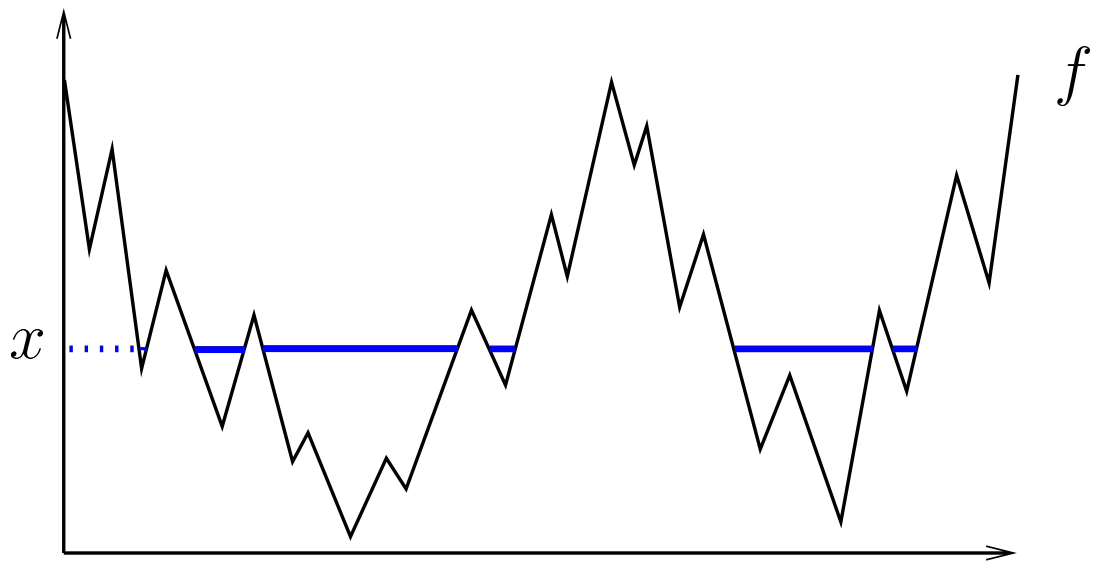
- *Topological persistence and simplification* by H. Edelsbrunner et. al. (2000)
- Topological approach for separating signal from noise.
- Data = real function over a topological space.

The idea of persistence



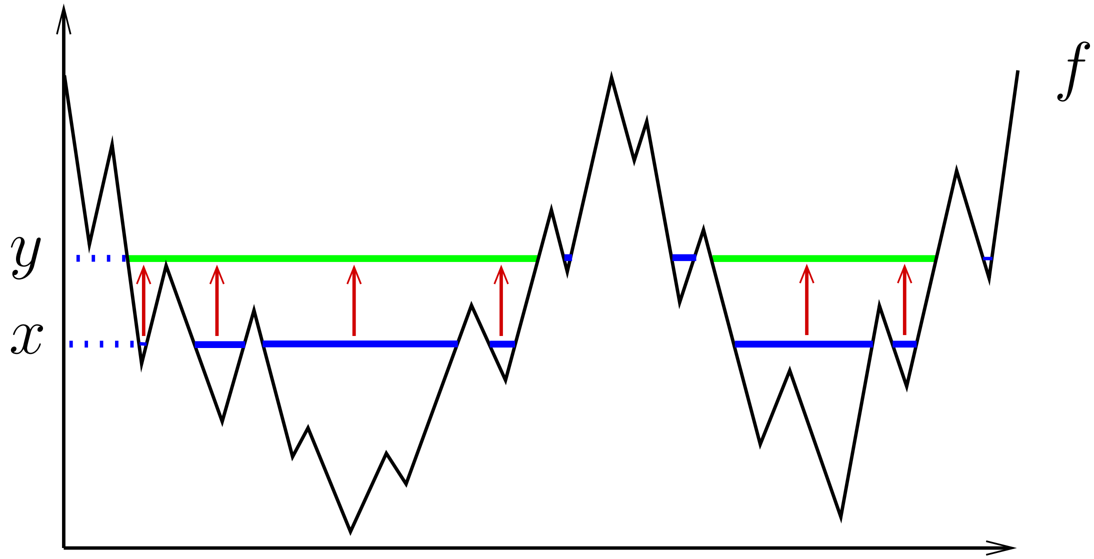
- How many components in $g^{-1}(-\infty, x]$?

The idea of persistence



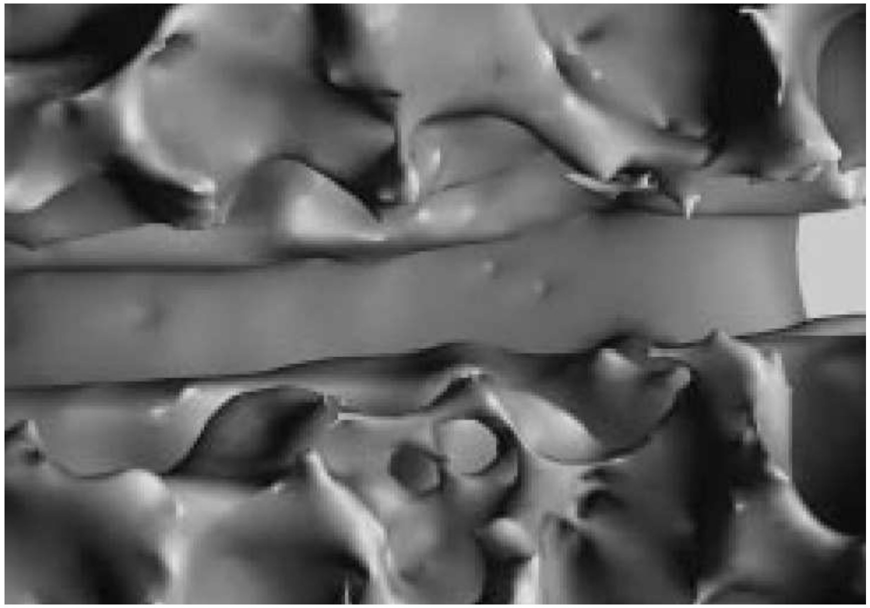
- How many components in $g^{-1}(-\infty, x]$?

The idea of persistence



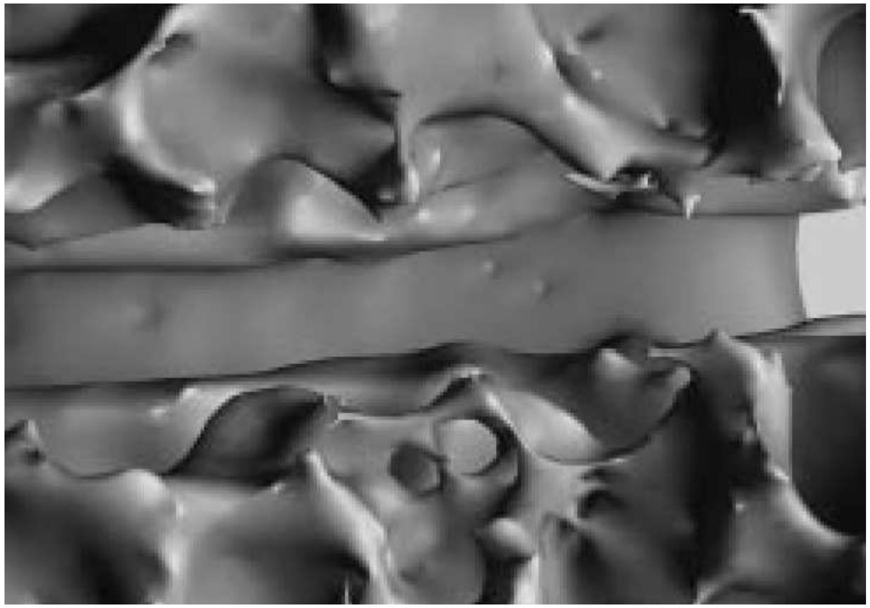
- How many components in $g^{-1}(-\infty, x]$?
- Count the components of $f^{-1}(-\infty, y]$ **induced** by those of $f^{-1}(-\infty, x]$.

Three-dimensional example



- What is the “actual” number of loops in this surface?
- More generally, how can we estimate the k-th **Betti number** of a sub-level set if the function is noisy?

Persistent Betti numbers

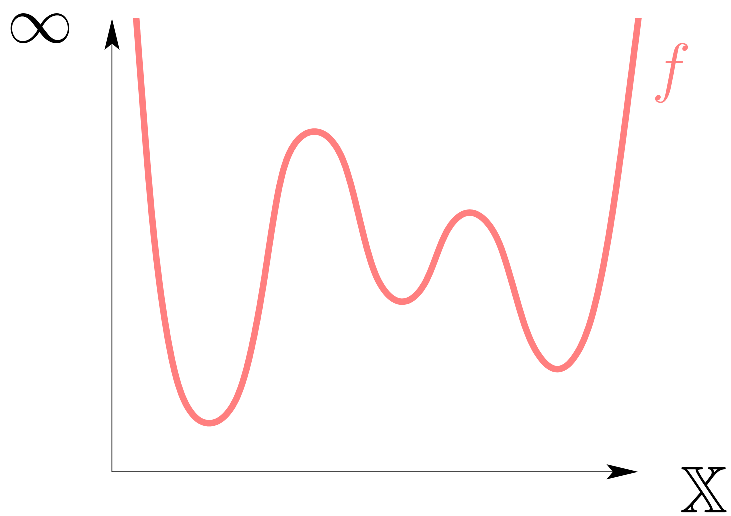


- **Persistent k-th Betti number** of $f : \mathbb{X} \to \mathbb{R}$:

$$\beta_k^{x,y}(f) = \mathrm{rk}(H_k(f^{-1}(-\infty, x]) \to H_k(f^{-1}(-\infty, y]))$$

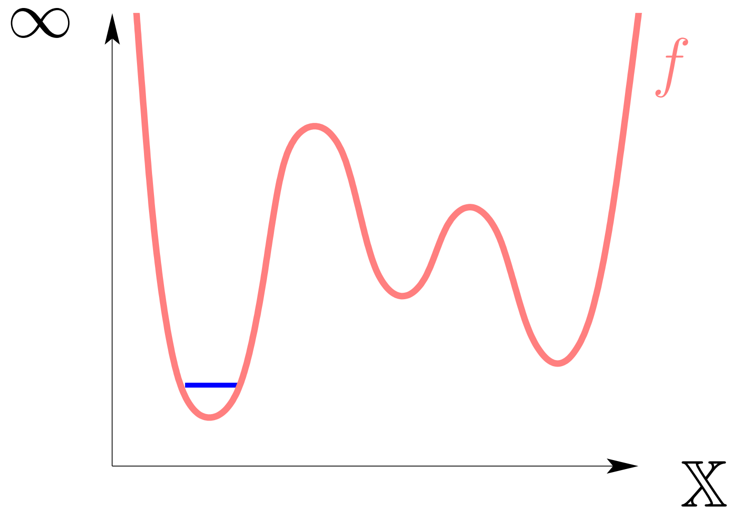
- Filter out **topological noise**.

Persistence intervals



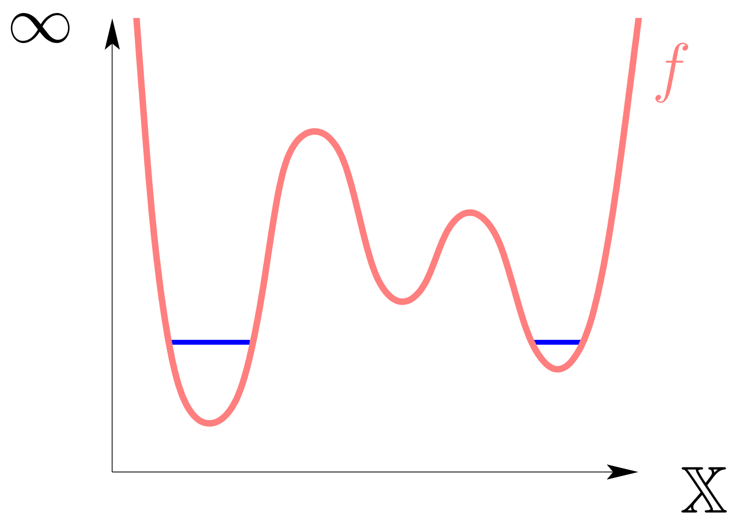
- Track the evolution of the topology of sub-level sets as the threshold increases.
- Pair thresholds that create components with those that destroy them.

Persistence intervals



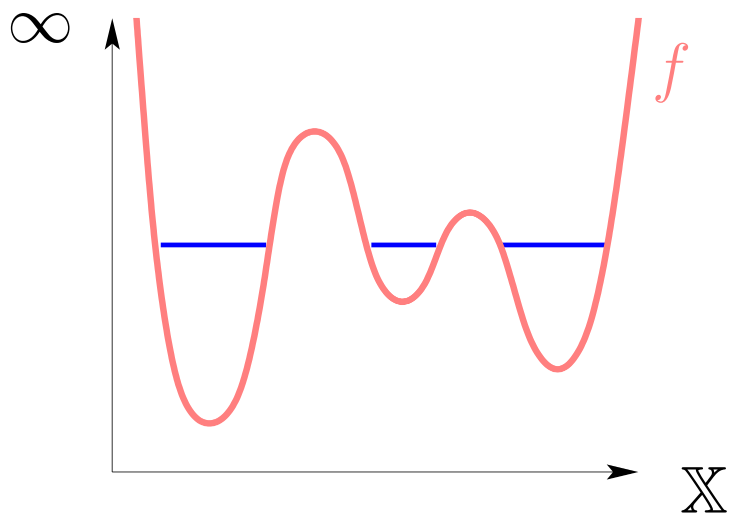
- Track the evolution of the topology of sub-level sets as the threshold increases.
- Pair thresholds that create components with those that destroy them.

Persistence intervals



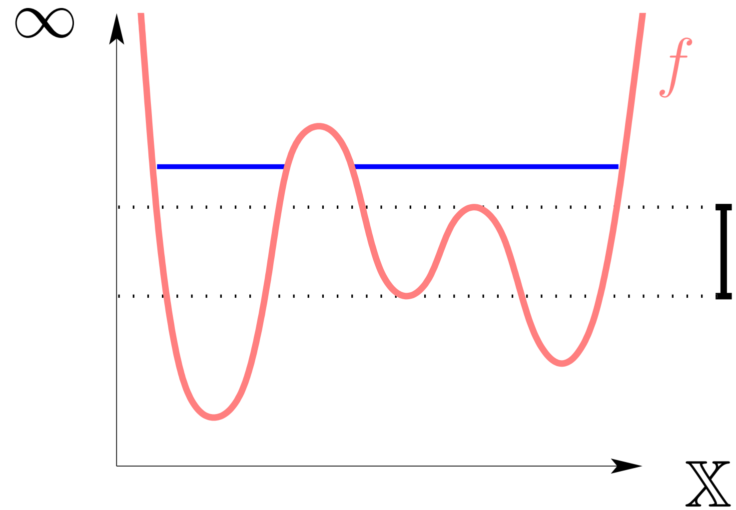
- Track the evolution of the topology of sub-level sets as the threshold increases.
- Pair thresholds that create components with those that destroy them.

Persistence intervals



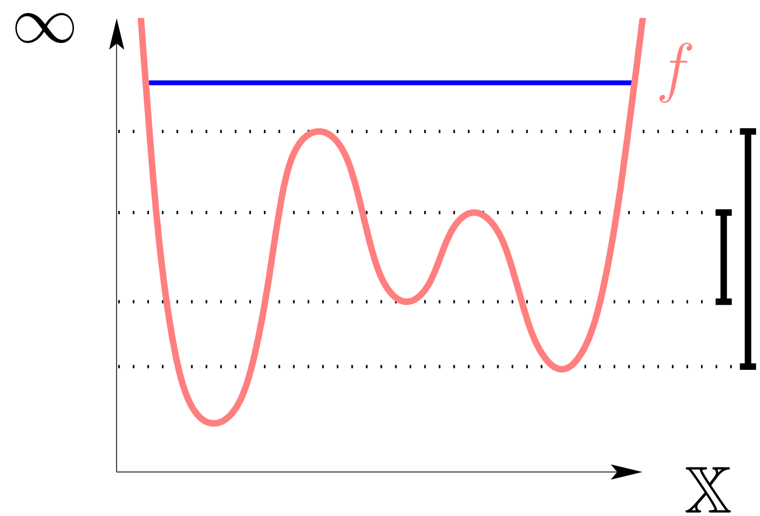
- Track the evolution of the topology of sub-level sets as the threshold increases.
- Pair thresholds that create components with those that destroy them.

Persistence intervals



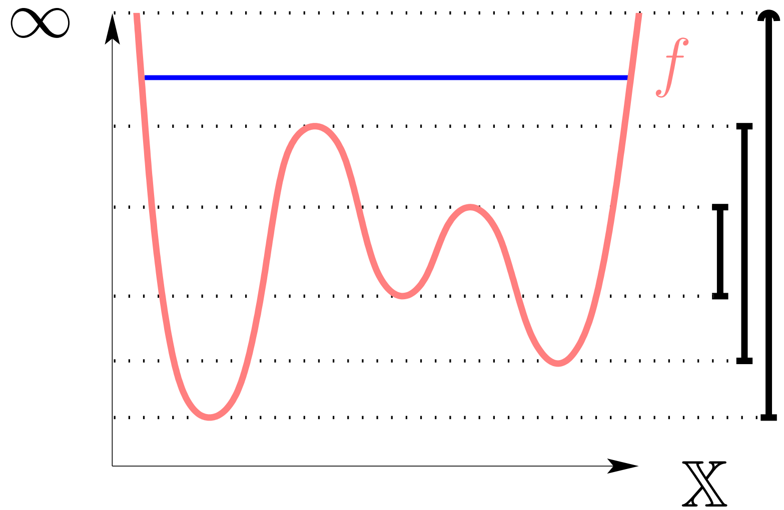
- Track the evolution of the topology of sub-level sets as the threshold increases.
- Pair thresholds that create components with those that destroy them.

Persistence intervals



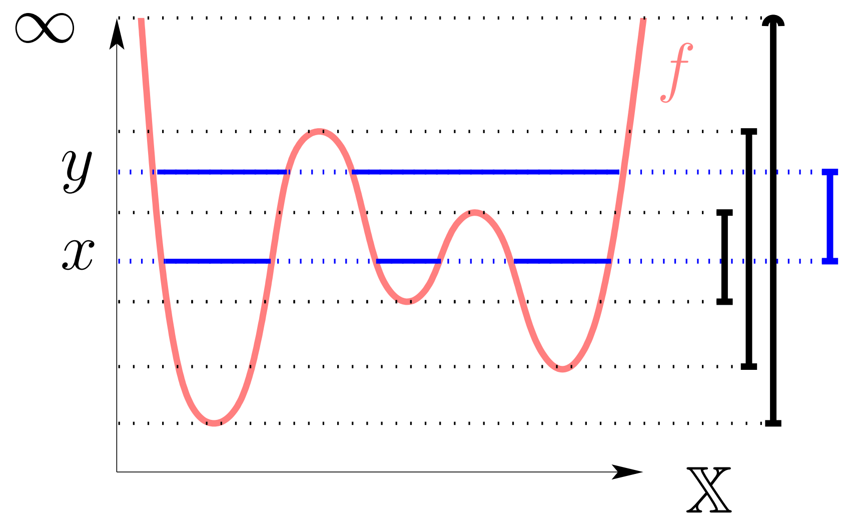
- Track the evolution of the topology of sub-level sets as the threshold increases.
- Pair thresholds that create components with those that destroy them.

Persistence intervals



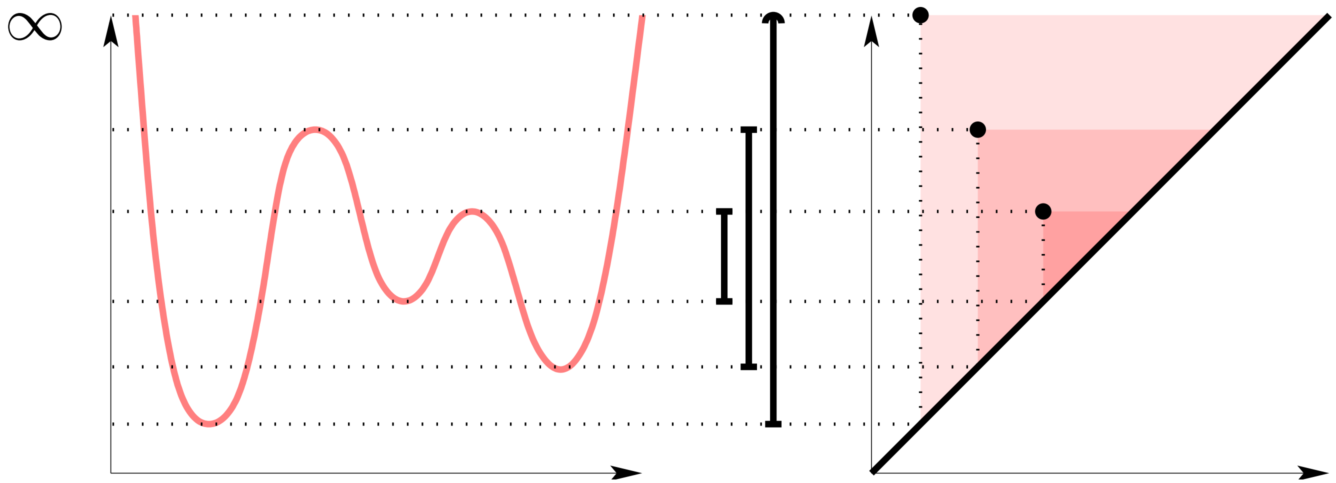
- Track the evolution of the topology of sub-level sets as the threshold increases.
- Pair thresholds that create components with those that destroy them.

Persistence intervals



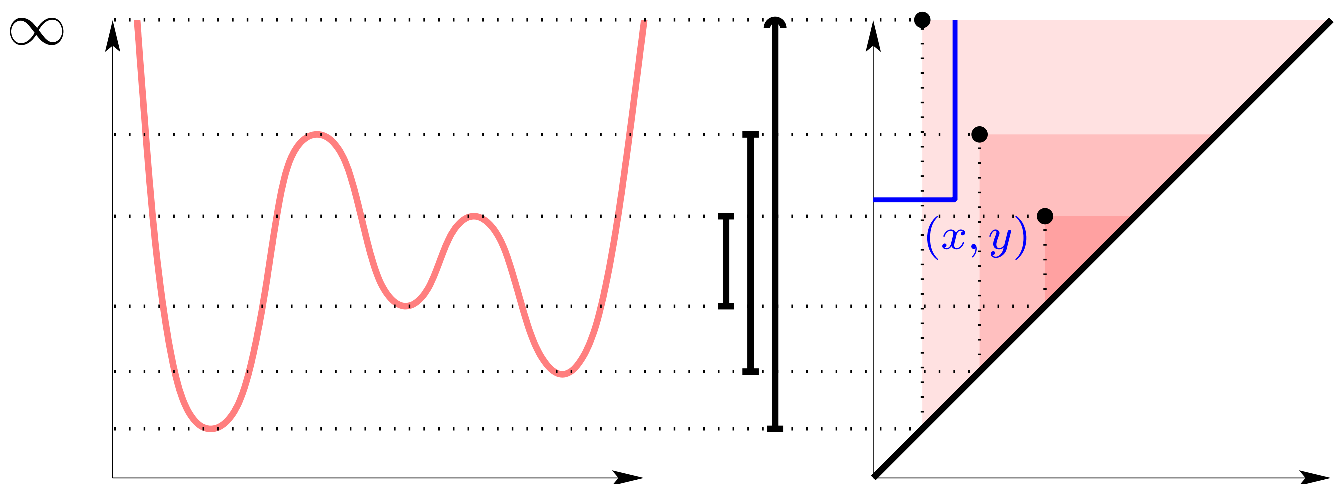
- The number of intervals containing $[x, y]$ is $\beta_k^{x,y}(f)$ ("k*-triangle lemma*" [ELZ02][CaZo04])
- In the case of components, persistence is equivalent to *Size theory*, developped by [Frosini, Landi].

Persistence diagrams



- Persistence intervals become points in the plane.
- The diagonal is included.

Persistence diagrams



- $\beta_k^{x,y}(f)$ is the number of points of $\mathrm{D}_k(f)$ within the upper left quadrant with corner (x, y).
- Persistence diagrams encode the topology of all sub-level sets at all *scales*.

Algorithm for PL functions

Persistence algorithm

Sort the simplices by increasing function values.

Build the mod 2 incidence matrix: $A_{ij} = 1$ iff $s_j \subsetneq s_i$.

while two columns have their last 1 on the same row

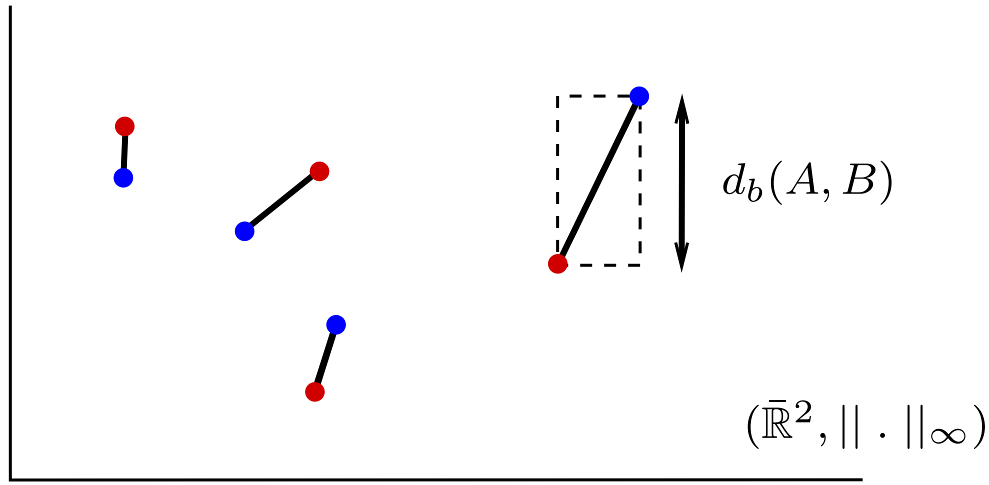
do

add the leftmost to the rightmost.

end while

return $\{(value(s_i), value(s_{last(i)})\}$

Metric on diagrams

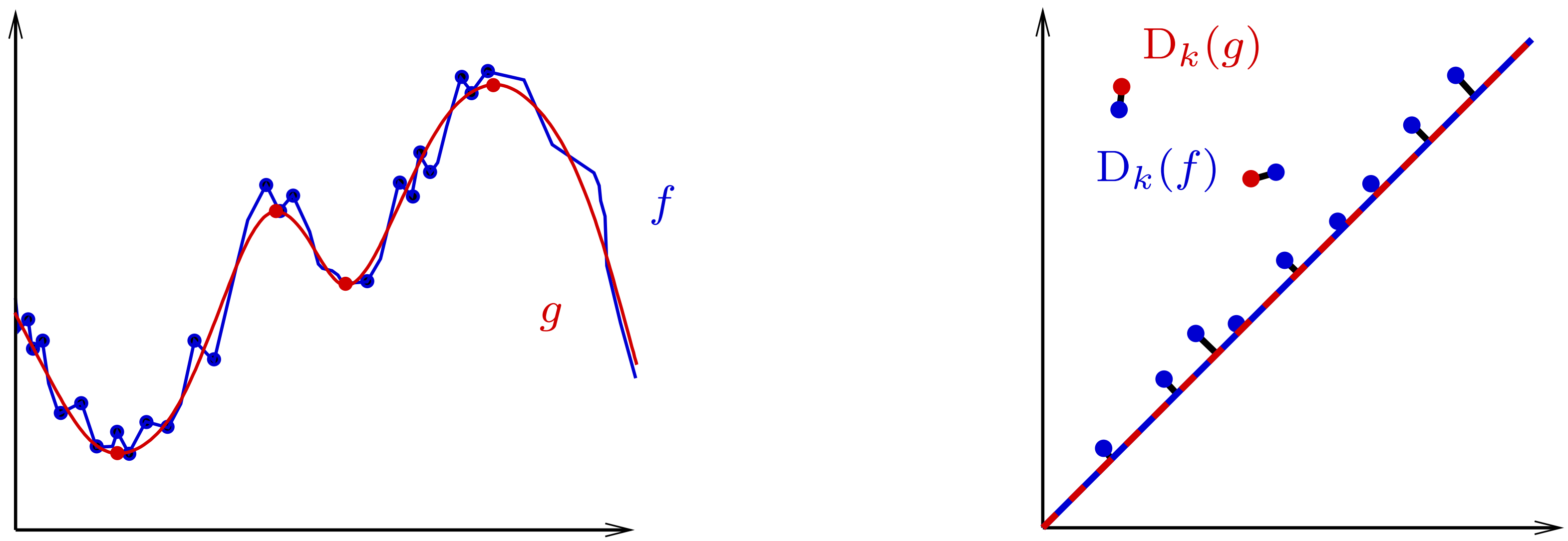


Definition. The *bottleneck distance* between sets A and B is:

$$d_b(A, B) = \inf_\gamma \sup_a \|a - \gamma(a)\|_\infty$$

over all $a \in A$ and all bijections $\gamma : A \to B$.

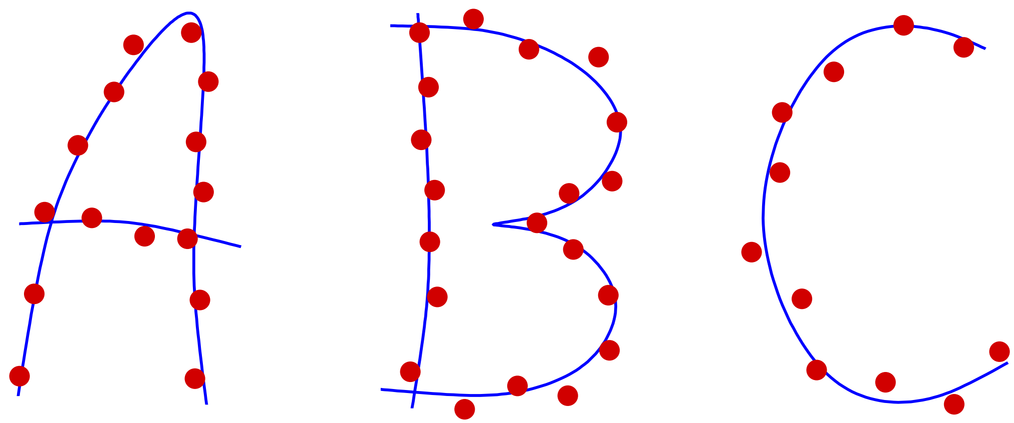
Stability theorem



Theorem. [CEH04].*For two continuous* tame *functions f and g on a finitely triangulable space:*

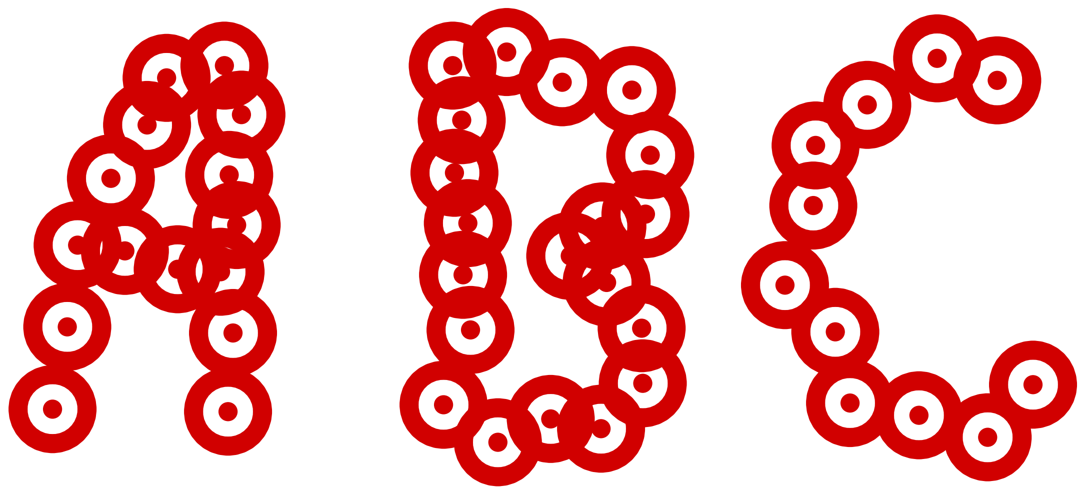
$$d_b(\mathrm{D}_k(f), \mathrm{D}_k(g)) \leq \|f - g\|_\infty$$

Betti numbers from samples



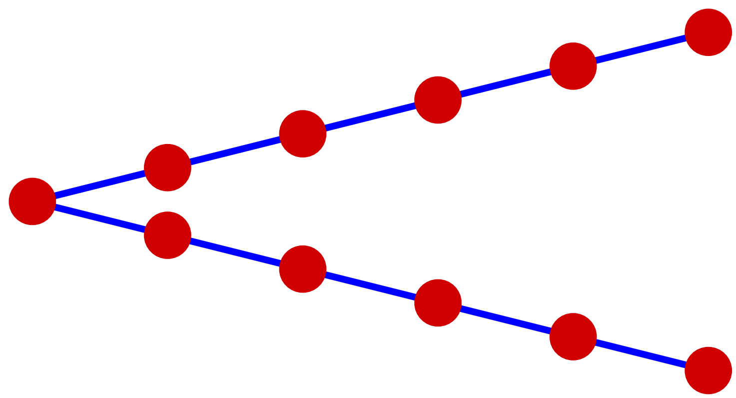
- Build a simplicial approximation of the unknown shape and compute its Betti numbers.
- Use offsets/alpha-shapes.

Reconstruction by offset

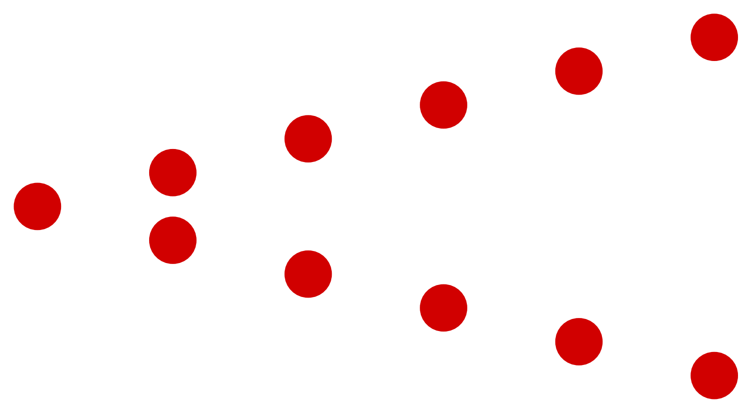


- Works for a large class of shapes in $\mathbb{R}^n$ [CCSL06].
- But might requires many data points.

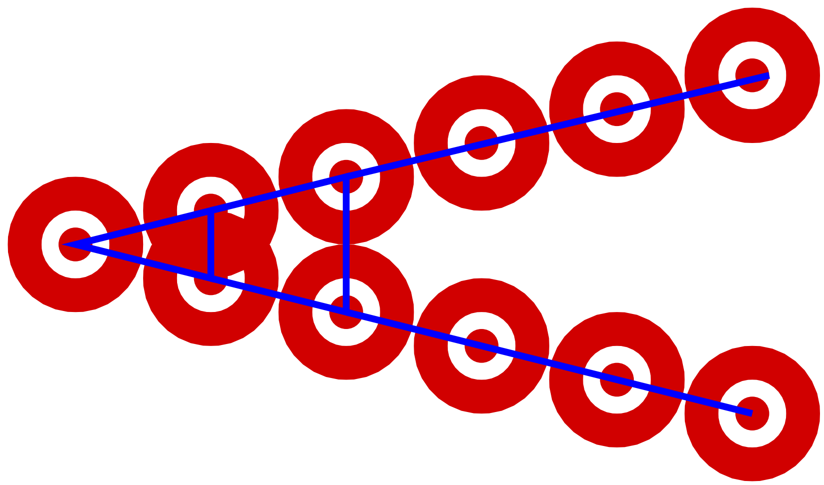
“Sharp” angle problem



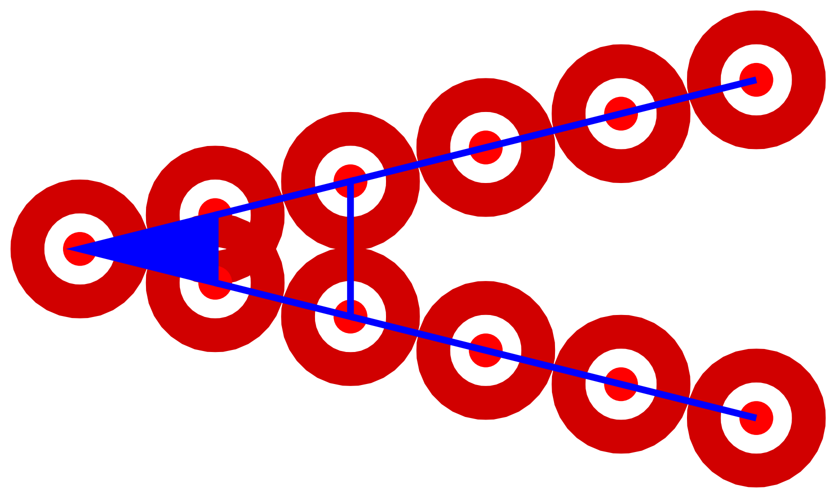
“Sharp” angle problem



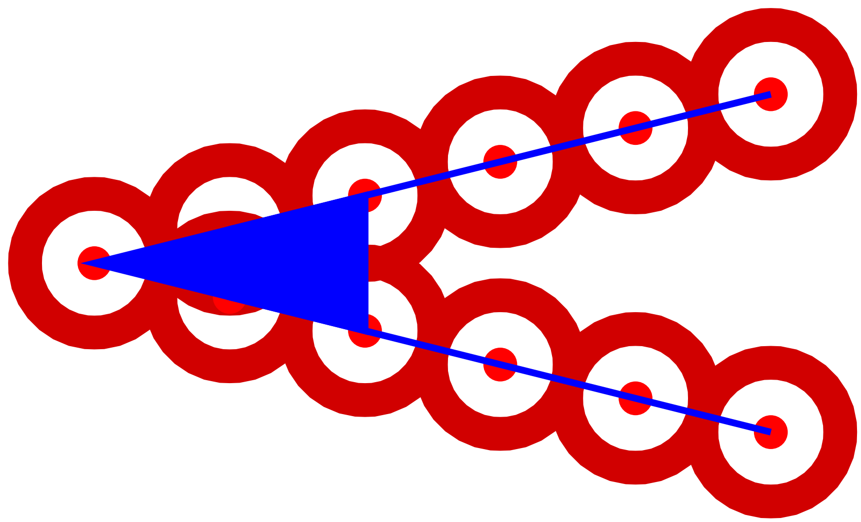
“Sharp” angle problem



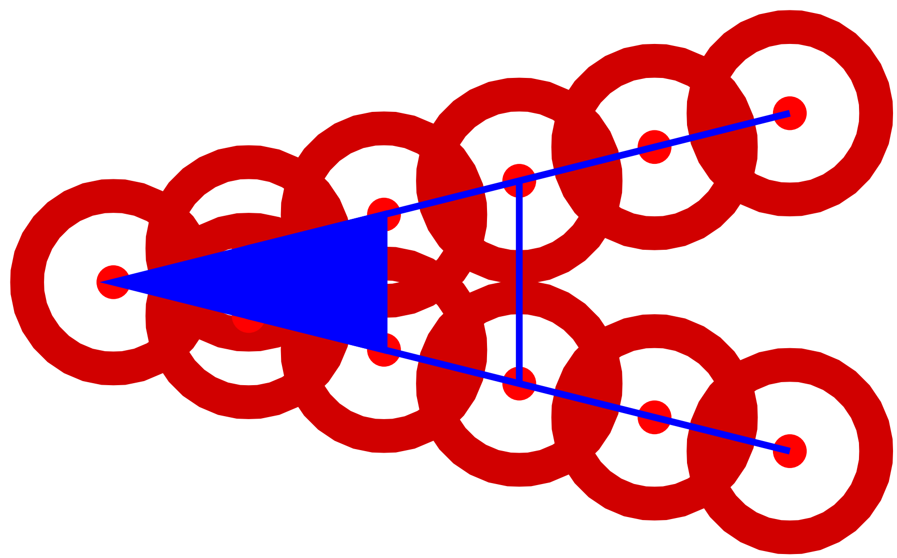
“Sharp” angle problem



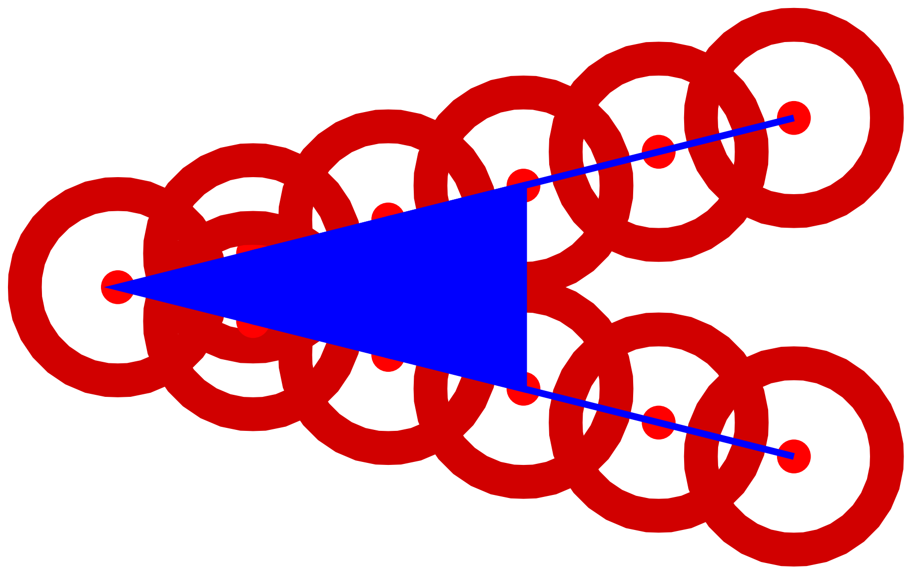
“Sharp” angle problem



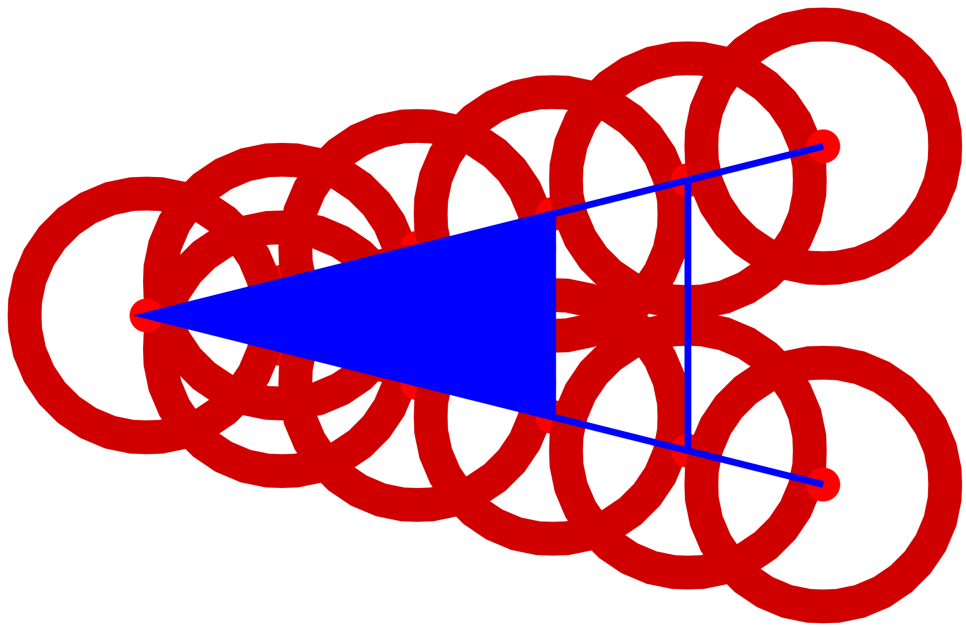
“Sharp” angle problem



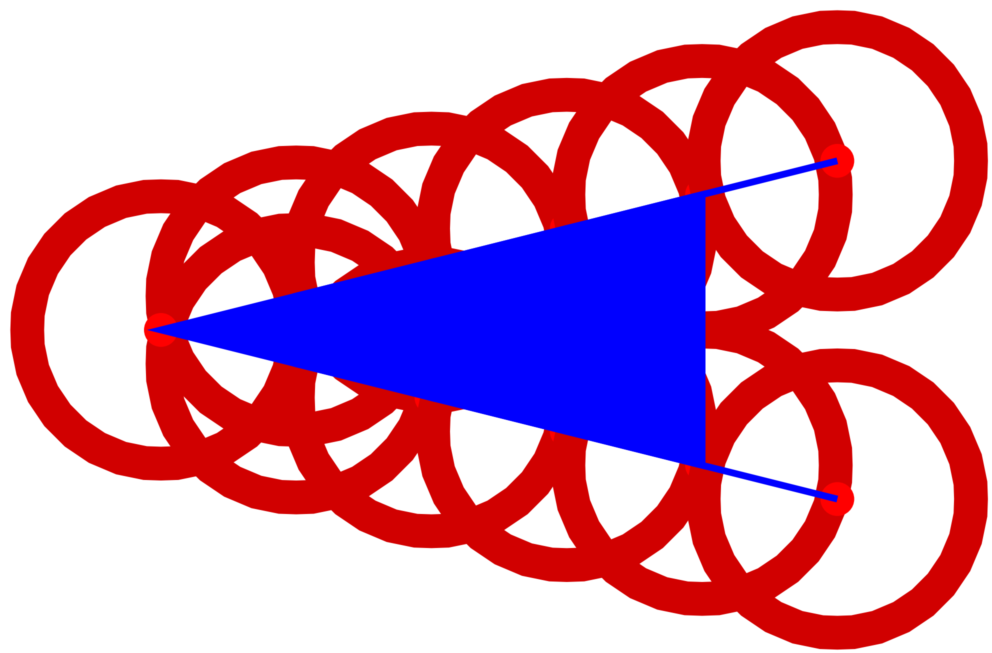
“Sharp” angle problem



“Sharp” angle problem



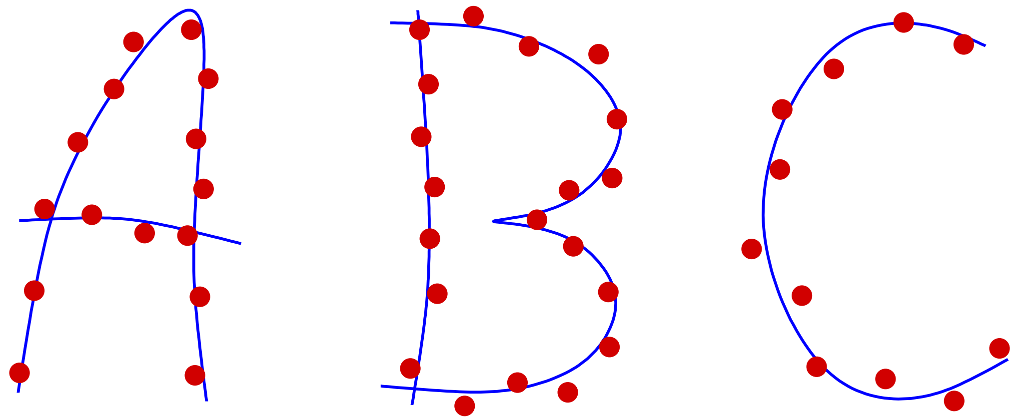
“Sharp” angle problem



Persistence helps

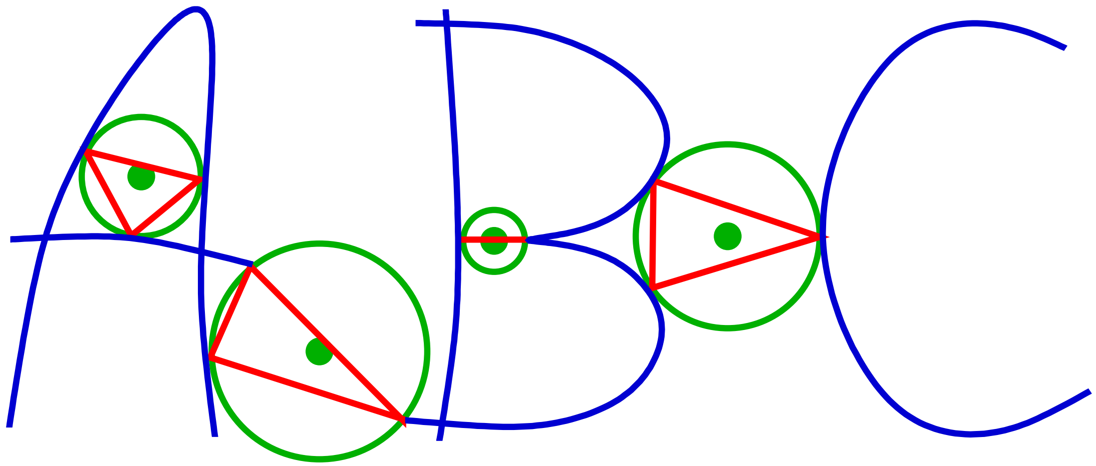
- *Distance function* to S:
  $\text{dist}_S(p) = \inf\{d(s,p) | s \in S\}$ for $p \in \mathbb{R}^n$.
- We were looking at the sublevel-sets of the distance function to the point cloud.
- Spurious loops are short-lived if the sampling is good enough.

Hausdorff distance



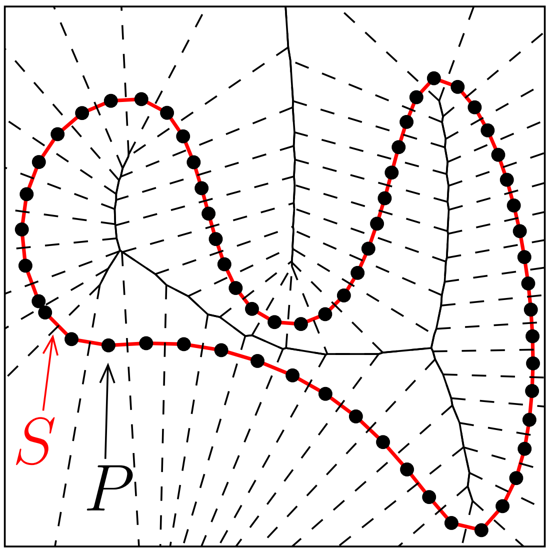
- The *Hausdorff distance* between two sets A and B is the minimum number r such that each point in A is at distance at most r from B and vice versa.
- $d_H(A, B) = ||\text{dist}_A - \text{dist}_B||_\infty$

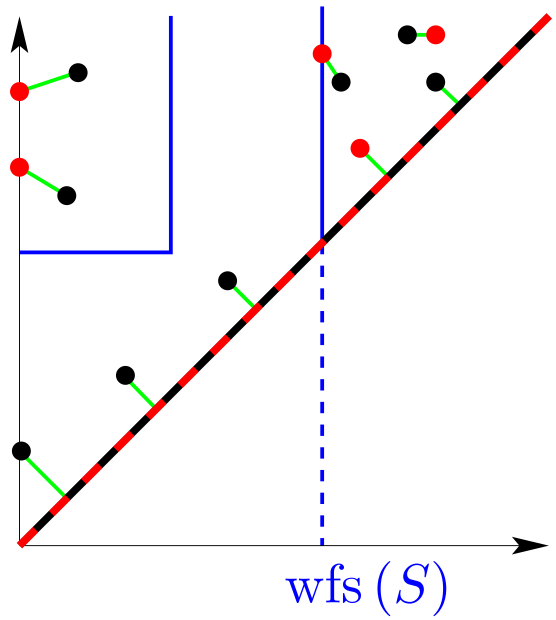
Weak feature size



- $\mathrm{wfs}\,(C) =$ inf $\{$ positive critical value of dist$_C\}$
- $\mathrm{wfs}\,(C) > 0$ if $C \subset \mathbb{R}^n$ is semi-algebraic [Fu95].

Betti numbers from samples





Theorem. [CEH/CL04]. Let S and P be closed subsets of $\mathbb{R}^n$.
If l is such that $d_H(S,P) < l < \mathrm{wfs}\,(S)/4$:

$$\beta_k(S) = \beta_k^{l,3l}(\text{dist}_P)$$

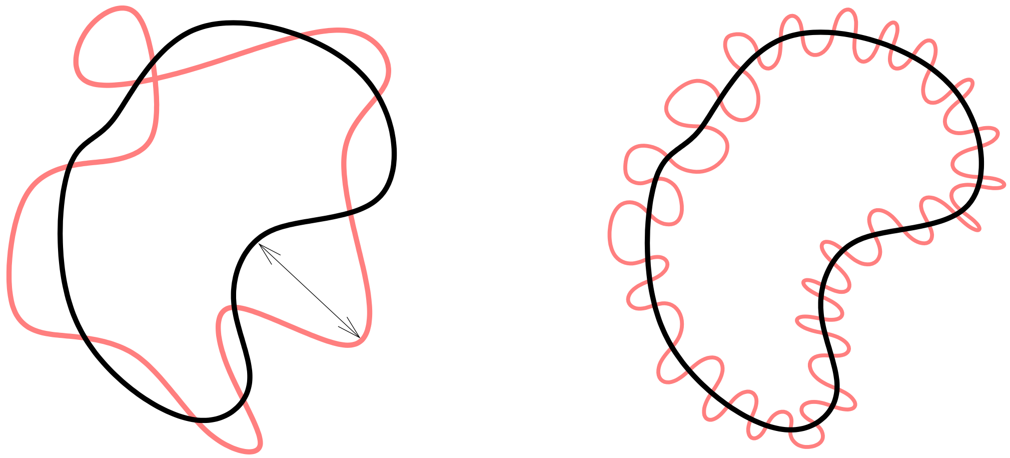
Comments

- Persistent Betti numbers/diagrams of distance function easily computable from the Delaunay triangulation of the sample points.
- You do not get any simplicial complex with the correct Betti numbers.
- Case of high dimensional ambiant space: witness complexes [CdS03]

Robust signatures of shapes

- Given two shapes, are they **approximately congruent**?
- Pick some rotation invariant function defined on shapes, *e.g.* distance function, curvature.
- Compare the persistence diagrams for the two shapes.

Problem for curves

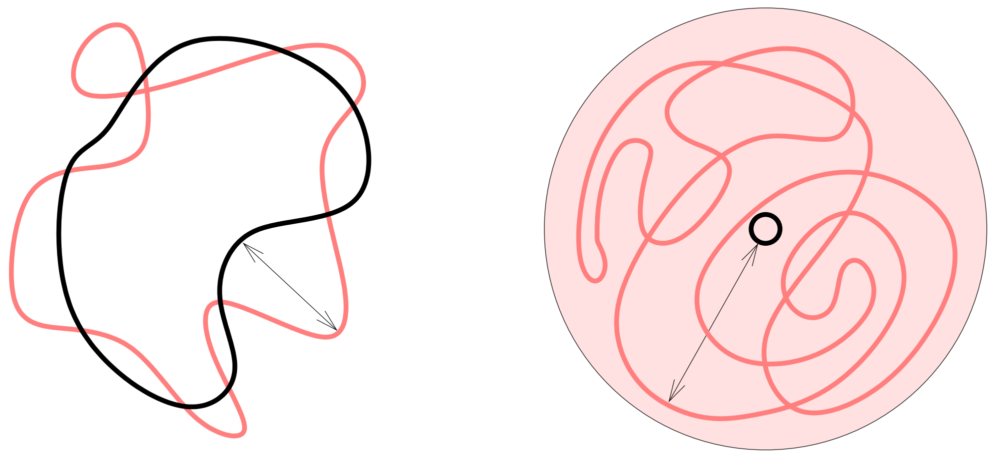


- If two curves are close, does it imply that their lengths are close?
- Fréchet distance between C_1 and C_2:

$$d_b(C_1, C_2) = \inf_{\phi_1, \phi_2} \sup_s d(\phi_1(s), \phi_2(s))$$

where ϕ_i ranges over all parameterizations of C_i.

Result



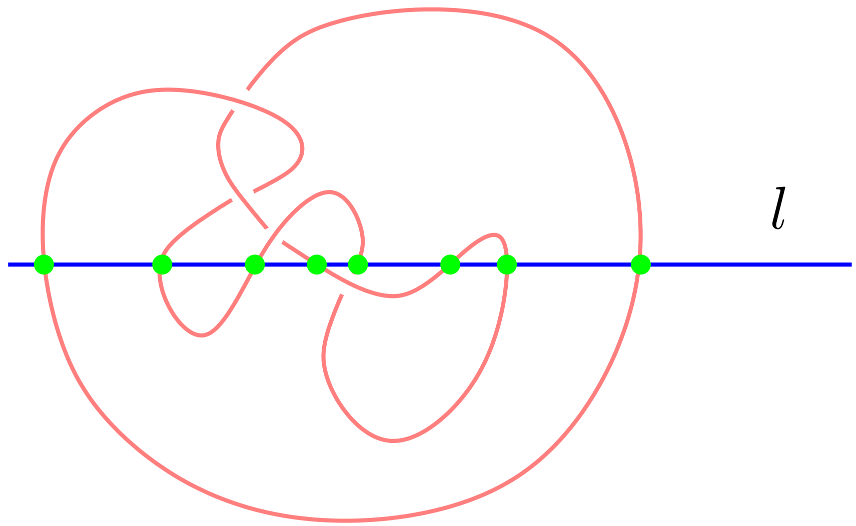
Theorem. *Let C_1 and C_2 be two closed curves in $\mathbb{R}^n$.*

Let L_i be the length of C_i, and K_i be the integral of its curvature.

One has:

$$|L_1 - L_2| \le \frac{2\mathrm{vol}(\mathbb{S}^{n-1})}{\mathrm{vol}(\mathbb{S}^n)}[K_1 + K_2 - 2\pi]\ d_b(C_1, C_2)$$

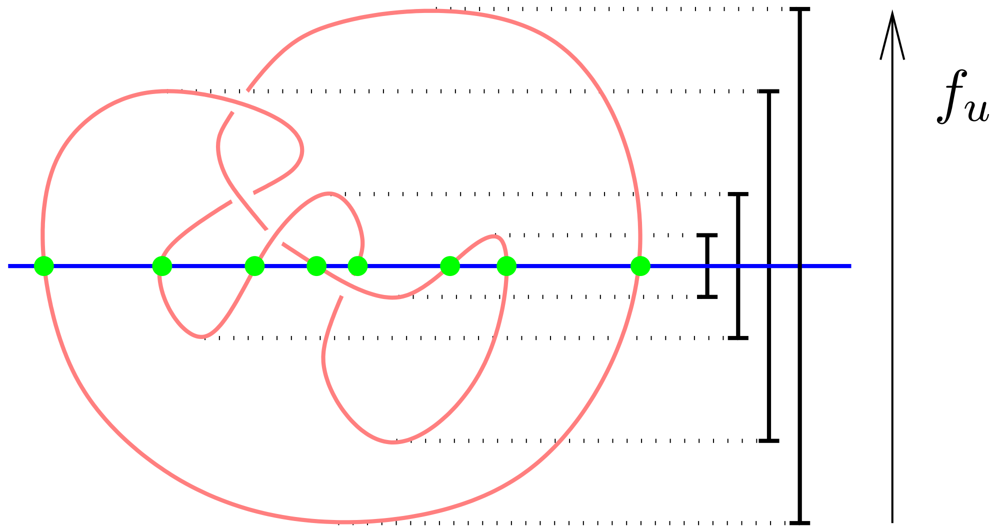
Proof



- Crofton formula:

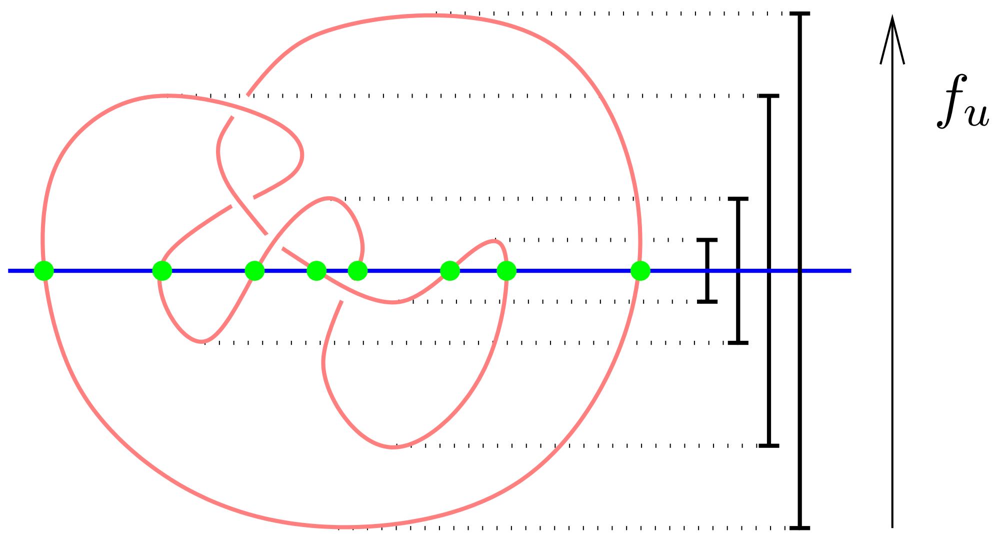
$$L(C) = \frac{\pi}{\mathrm{vol}(\mathbb{S}^n)} \int_{\text{hyperplane } l \subset \mathbb{R}^n} \sharp(l \cap C)$$

Proof



- Let $f^u : C \to \mathbb{R}$ be the height function in the direction u.
- If l has normal vector u, then $\sharp(l \cap C)$ is twice the number of "persistence intervals" of f_u stabbed by l.

Proof

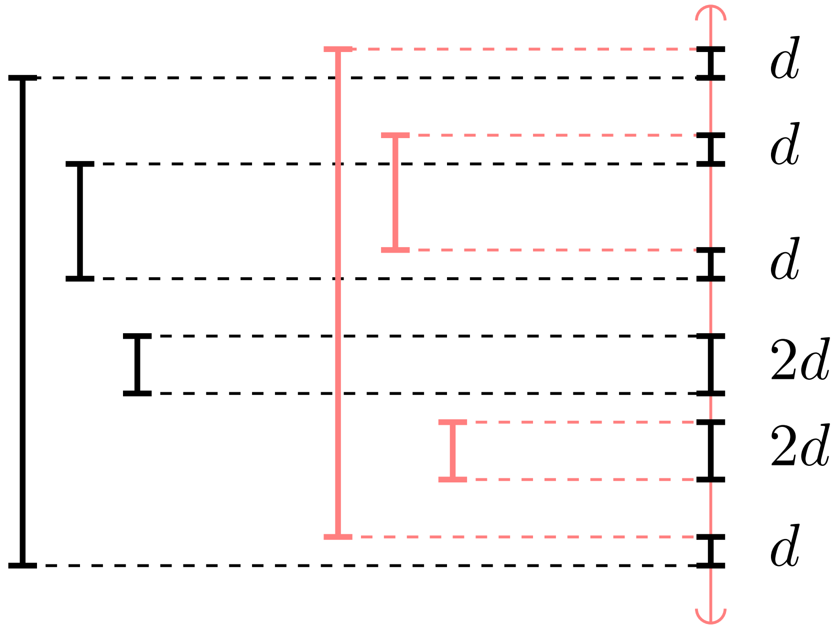


- Hence :

$$\int_{l \text{ hyperplane with normal } u} \sharp(l \cap C)$$

is twice the total length of the persistence intervals of f^u.

Proof



- Stability theorem : the bounds of the persistence intervals of f^u move by at most $d_b(C_1, C_2) = d$.
- Hence the total length of these intervals changes by at most $d(n_1^u + n_2^u - 2)$, where n^u is the number of critical points of f^u.

Proof

- By integrating over all directions :

$$|L_1 - L_2| \leq 2d \, \frac{\pi}{\mathrm{vol}(\mathbb{S}^n)} \int_{u \in \mathbb{S}^n} n_1^u + n_2^u - 2 \, du$$

- *Exchange theorem*:
  The integral of the number of critical points n_i^u over $u \in \mathbb{S}^n$
  is the integral of the curvature of C_i divided par $\pi/\mathrm{vol}(\mathbb{S}^{n-1})$
  $\rightarrow$ qed.

Result for surfaces

Theorem. *Let $S_1 = \partial V_1$ and $S_2 = \partial V_2$ be two closed surfaces in $\mathbb{R}^3$ with the same genus g. Let H_i be the integral of the* ***mean curvature*** *of S_i, and K_i be the integral of its* ***absolute Gauss curvature****. One has:*

$$|H_1 - H_2| \leq [K_1 + K_2 - 4\pi(1+g)]\, d_b(V_1, V_2)$$

- Holds for piecewise-linear surfaces, for which simple formula exist: accurate total mean curvature estimation from a mesh.
- Closeness between normals to the surfaces is not explicitly required, unlike in [CSM03].

Conclusion

- Thank you!